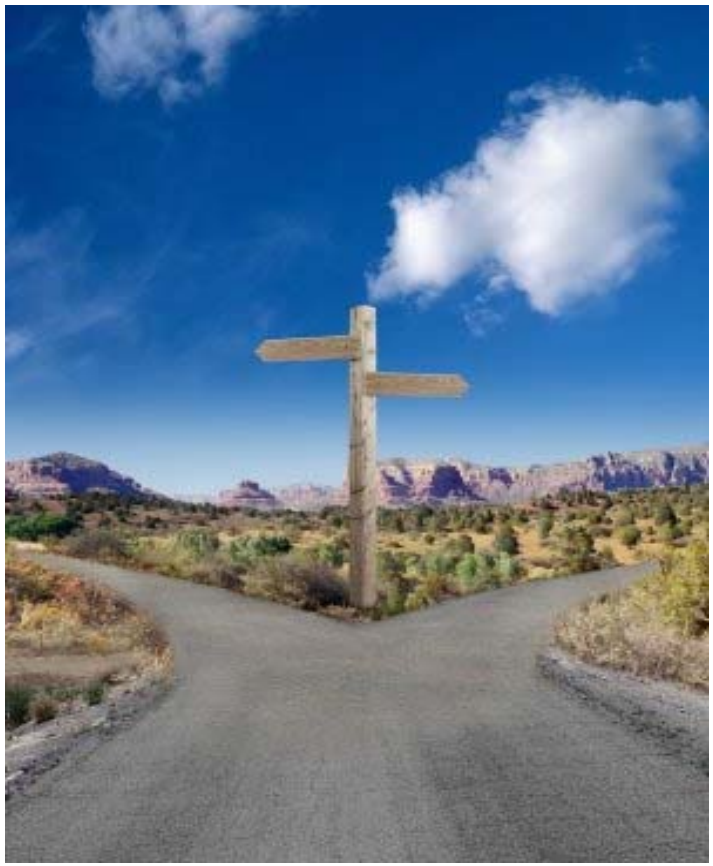
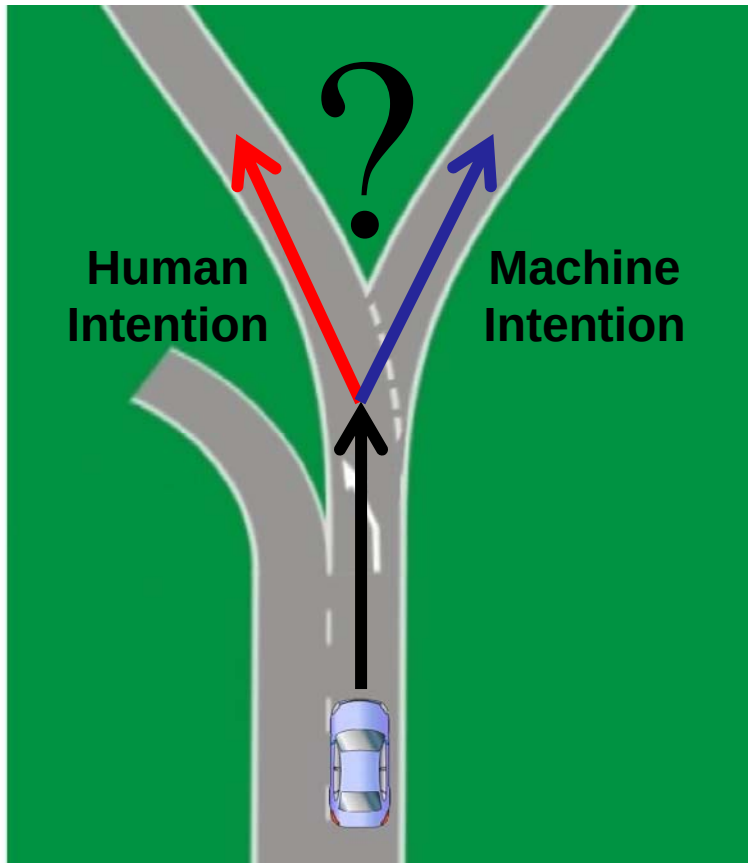
Arbitration between Driver and Automation:
why overriding is just the tip of the iceberg

Johann Kelsch

German Aerospace Center, Institute of Transportation Systems

Let‘s imagine driving down a road fork by an intelligent car



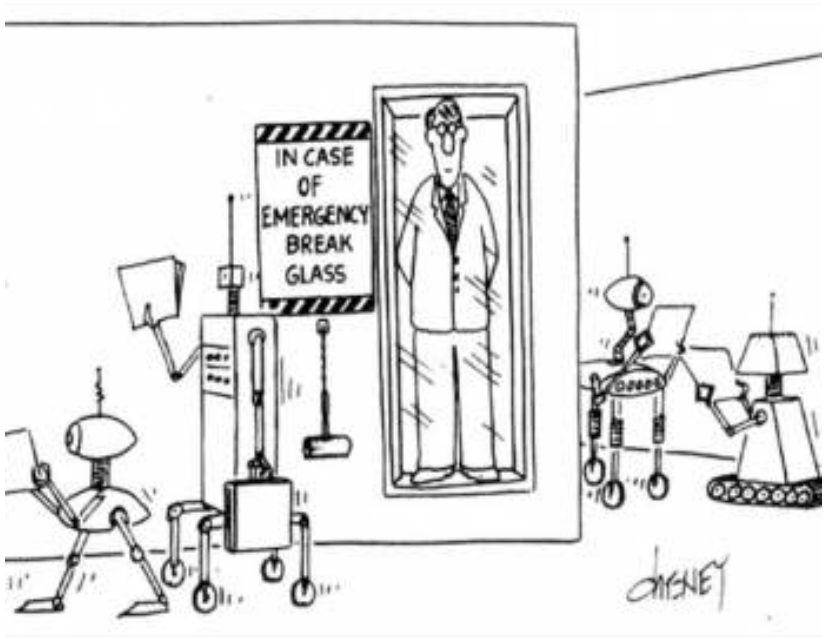


What will happen if human and machine have different intentions?

Picture source: 37stories.wordpress.com 2012

Necessity of overriding strategies

- **On the one hand:** “Every driver shall at all times be able to **control** his vehicle or to **guide** his animals.“ (Vienna Convention 1968)
- **On the other hand:** Growing technical possibilities to automate (parts of) the vehicle guidance (better sensors and actors, more computing power etc. 2012)



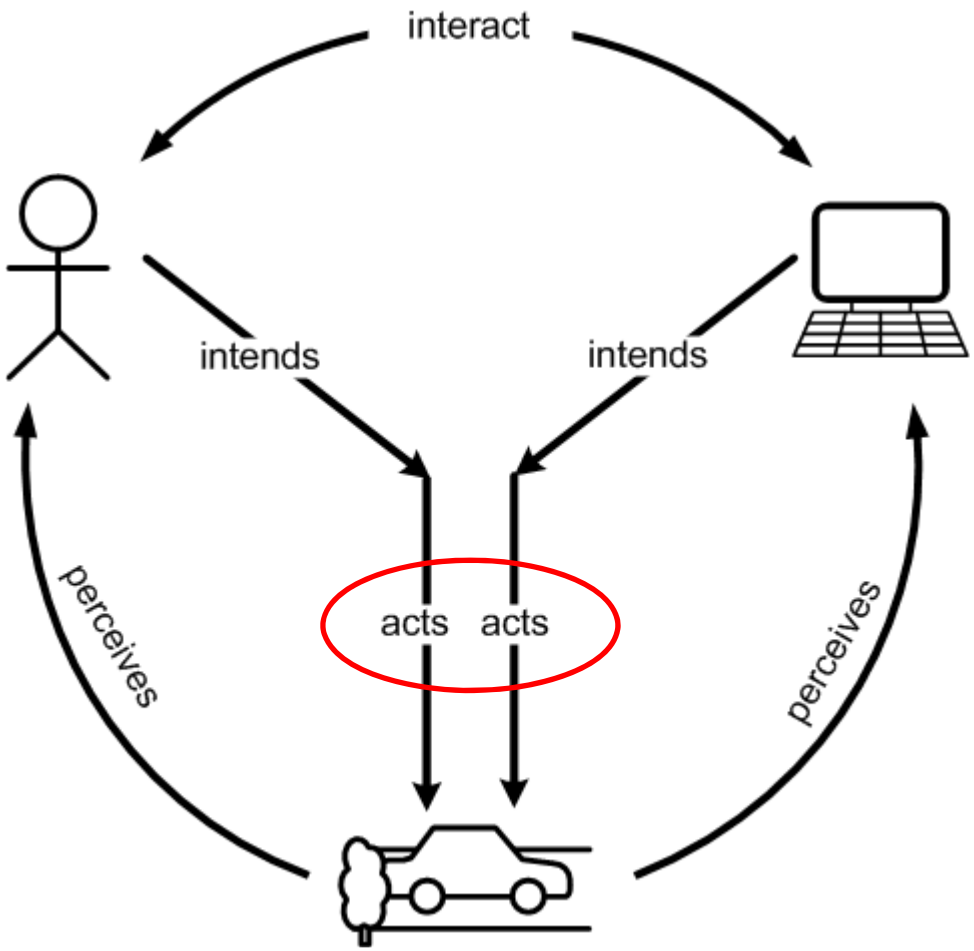
- Implications in **traffic law**
  - Driver/automation role/authority discussion
  - OEM responsibility discussion
  - ‘Translation constructions’ (control = monitoring) etc.
- Implications in **human – machine system design** for automotive
  - E.g. imperative of **overriding** strategies (comic)

Picture source: Cheney, 1989, News Yorker Magazine, Inc.

Overview

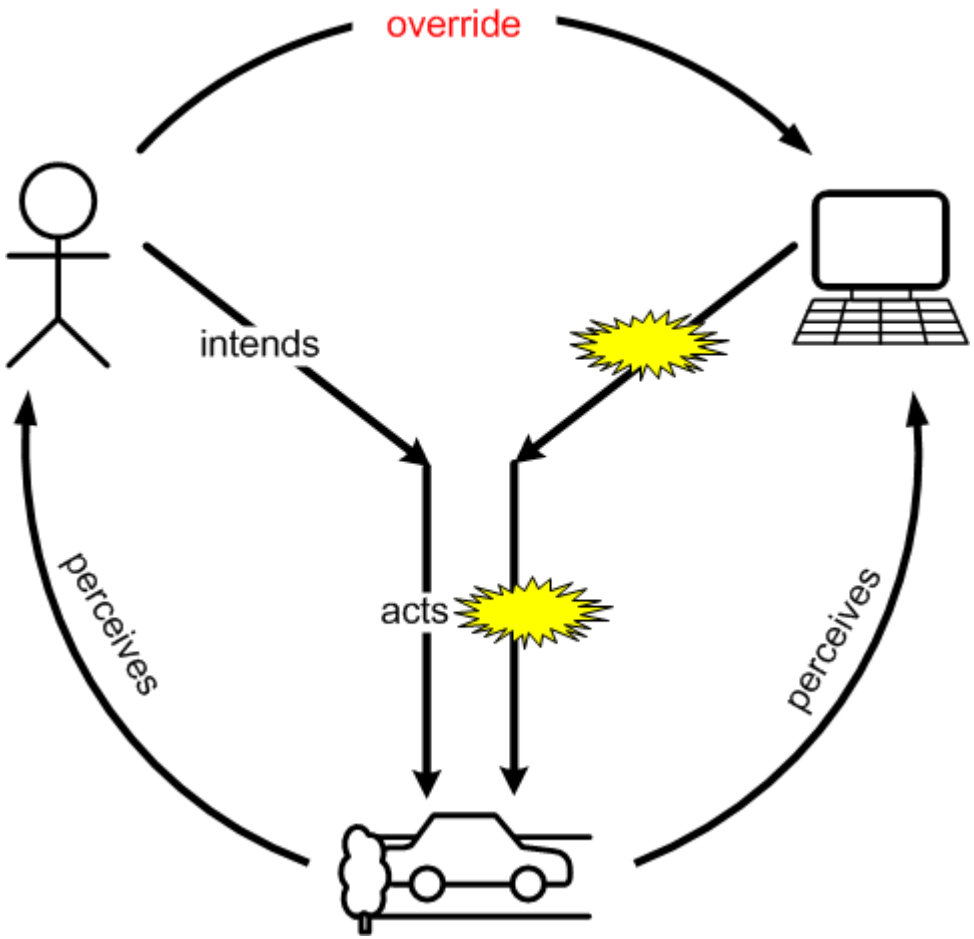
- Introduction: from **overriding** to **arbitration**
- Description of the human-machine arbitration **concept**
- Arbitration **design patterns** and **tools**
- **Empirical studies** in respect to InteractIVe
- Outlook

What shall we override at all?



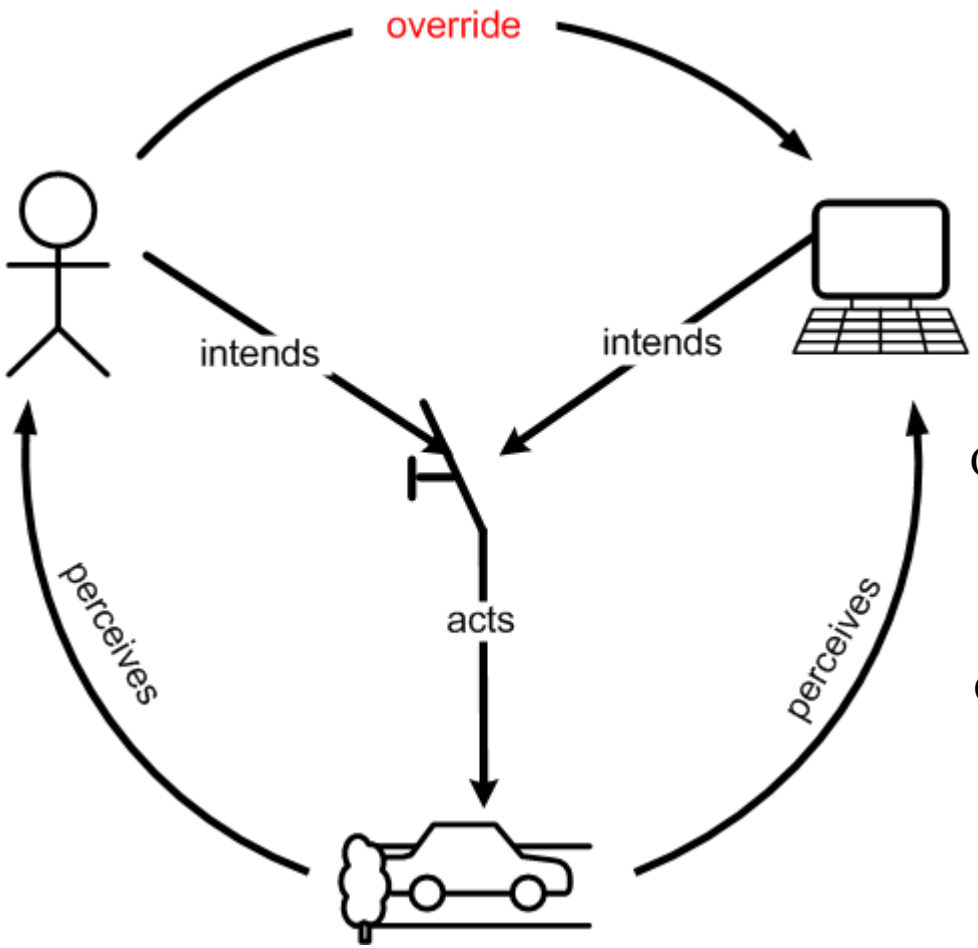
- Simplified **human** perception-action model
- **Machine** as a **cognitive agent** Hollnagel & Woods, (1983)
- Human and machine **interacts** with each other
- Human and machine **compete** for vehicle control
- **But:** human should be able to have the control exclusively!

What shall we override at all?

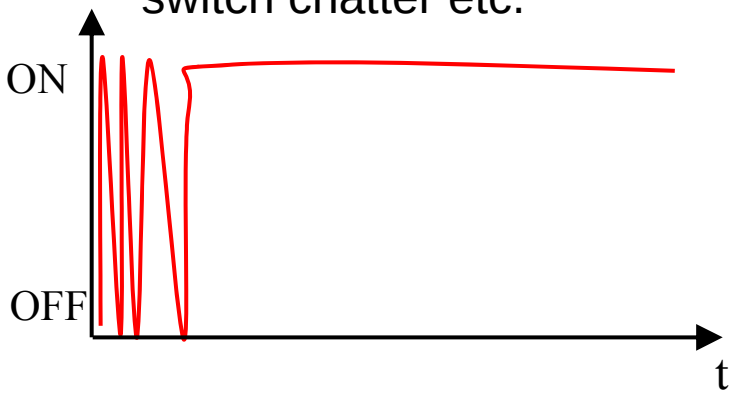


- Therefore, **overriding** of the machine action is necessary
- For consistent system **behavior design** we shall override machine intents too

Overriding as an ‘intent-act switch’

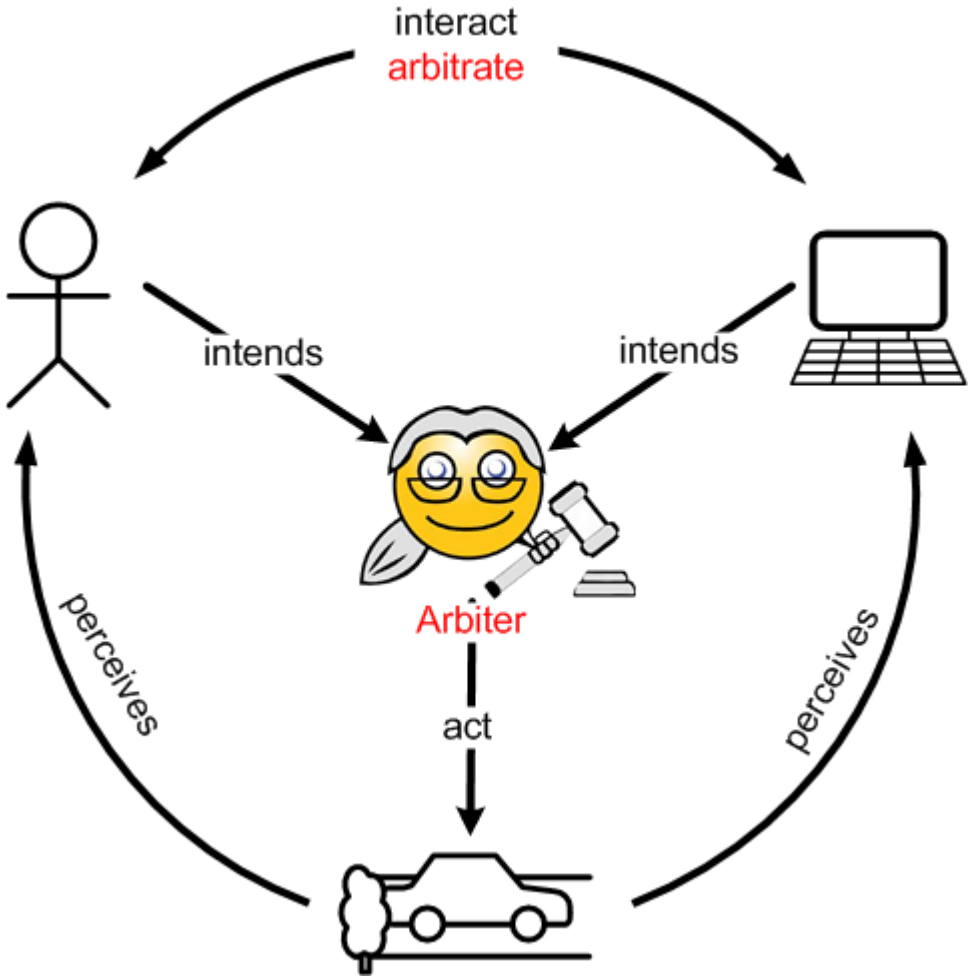


- Very simple kind of **overriding** implementation
- **Instability** problems with a ‘simple switch’ like switch chatter etc.



- **Similar** instability problems with a sudden switch between human and machine control intention

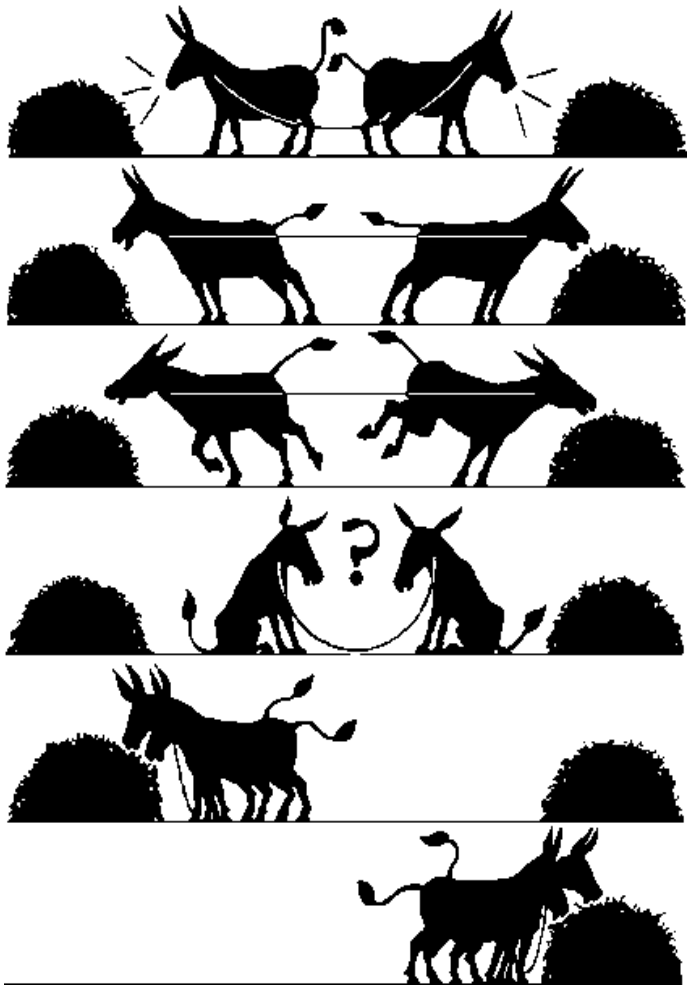
Overriding as a special form of arbitration



- Solution may be an ‘intelligent switch’ called **arbiter** and the **arbitration process**
- Technically: Arbiters are electronic modules that allocate access to **shared resources** (e.g. communication **deconflicting** in CAN-Bus)
- **Concept of arbitration** is also useful in human-machine **interaction design**

Smiley clipart source: clker.com edited

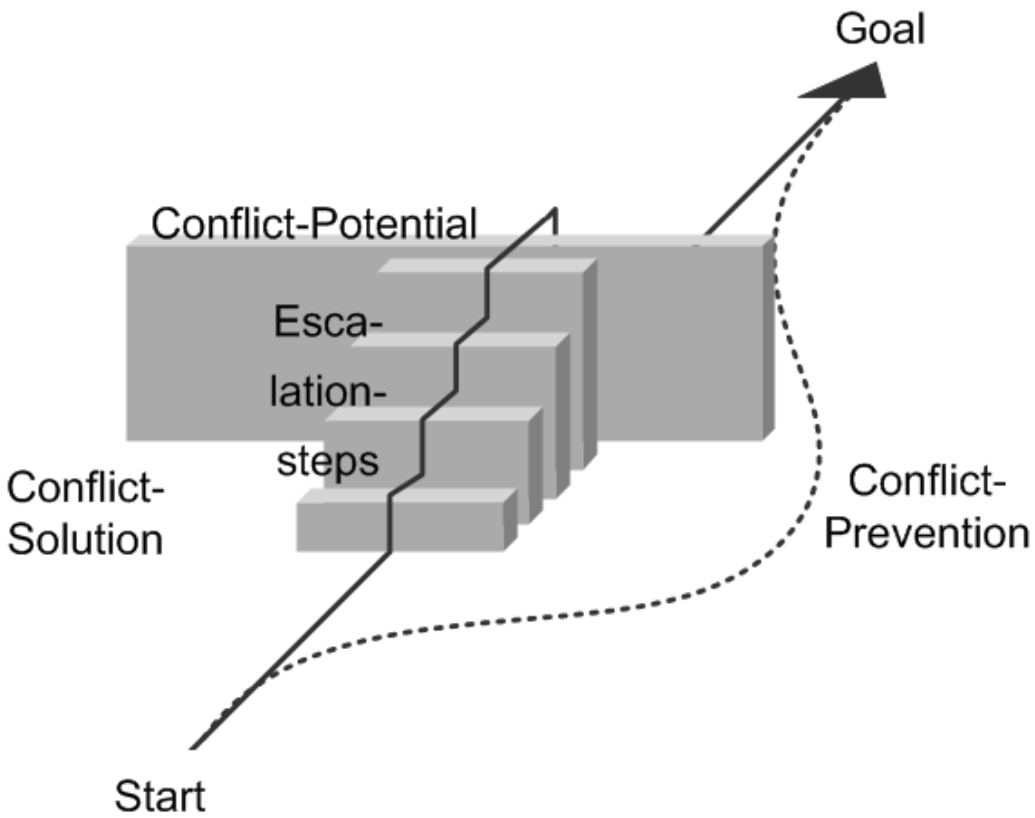
What is human - machine arbitration?



Human – machine **arbitration** we call a **finite negotiation** by means of proper **interaction** strategies aimed to reach a **joint intent** and an **adequate action** of the human-machine system within the **available time**.

Picture source: http://www.esci.at/eusipo/as3.pdf edited

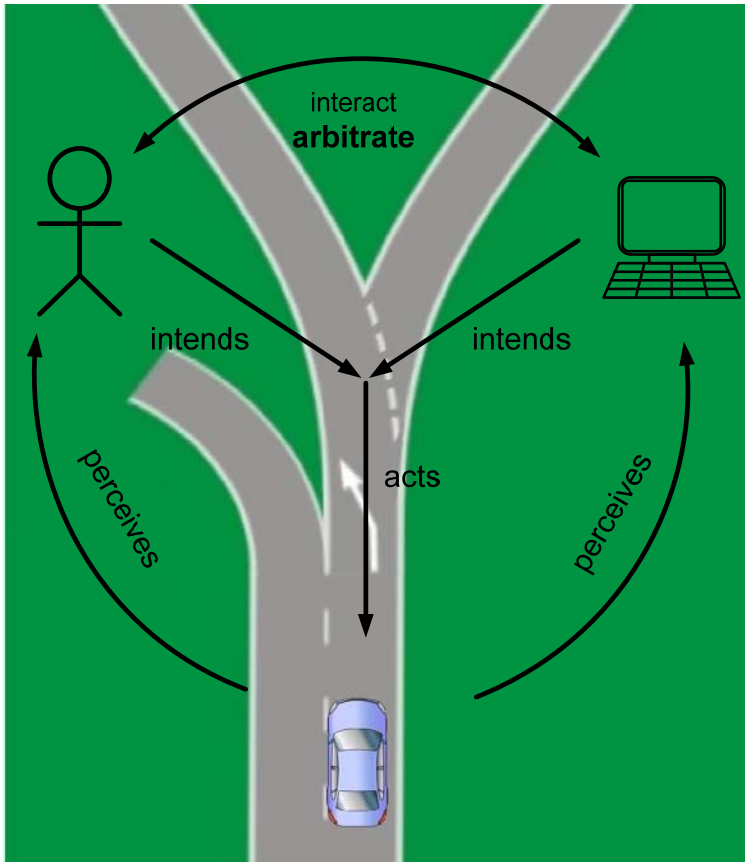
Additional remarks about arbitration



- Arbitration is about **system design**
- Arbitration deals with **conflicts**
- With the help of arbitration strategies one can **solve or prevent (*)** conflicts
- Arbitration needs an **analysis** for recognition of **conflict potential** within the designed system
- **Escalation** steps are **one of possible** arbitration strategies

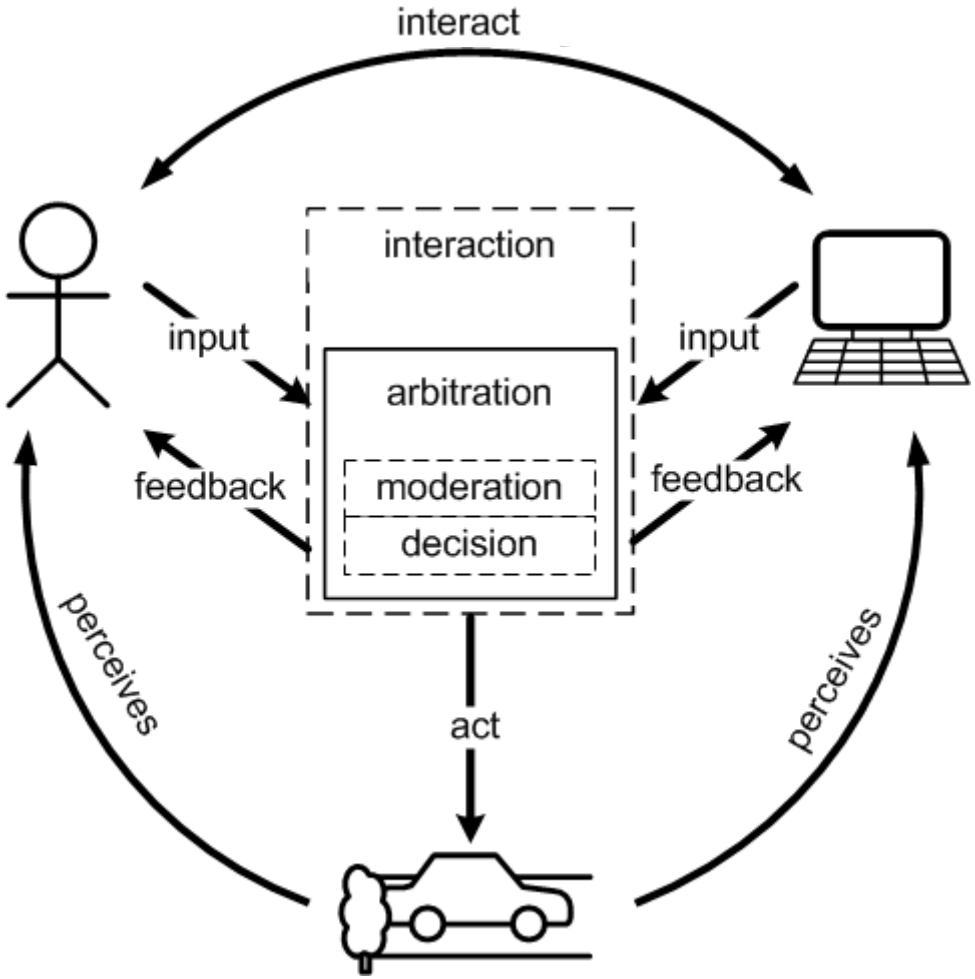
(*) after Griesche et al. 2012

What is the most critical arbitration in automotive?

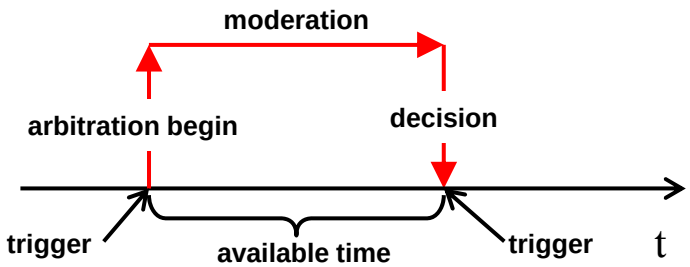


Driver – automation **control arbitration** we call a finite negotiation by means of proper interaction strategies aimed to reach a **joint control intent** and an adequate **control action** of the driver - automation system within the available time.

Arbitration time perspective: Sequence of Arbitration

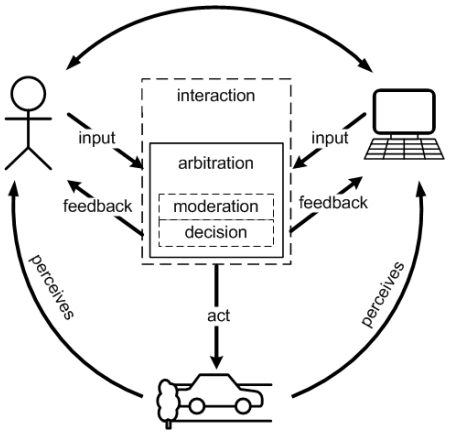


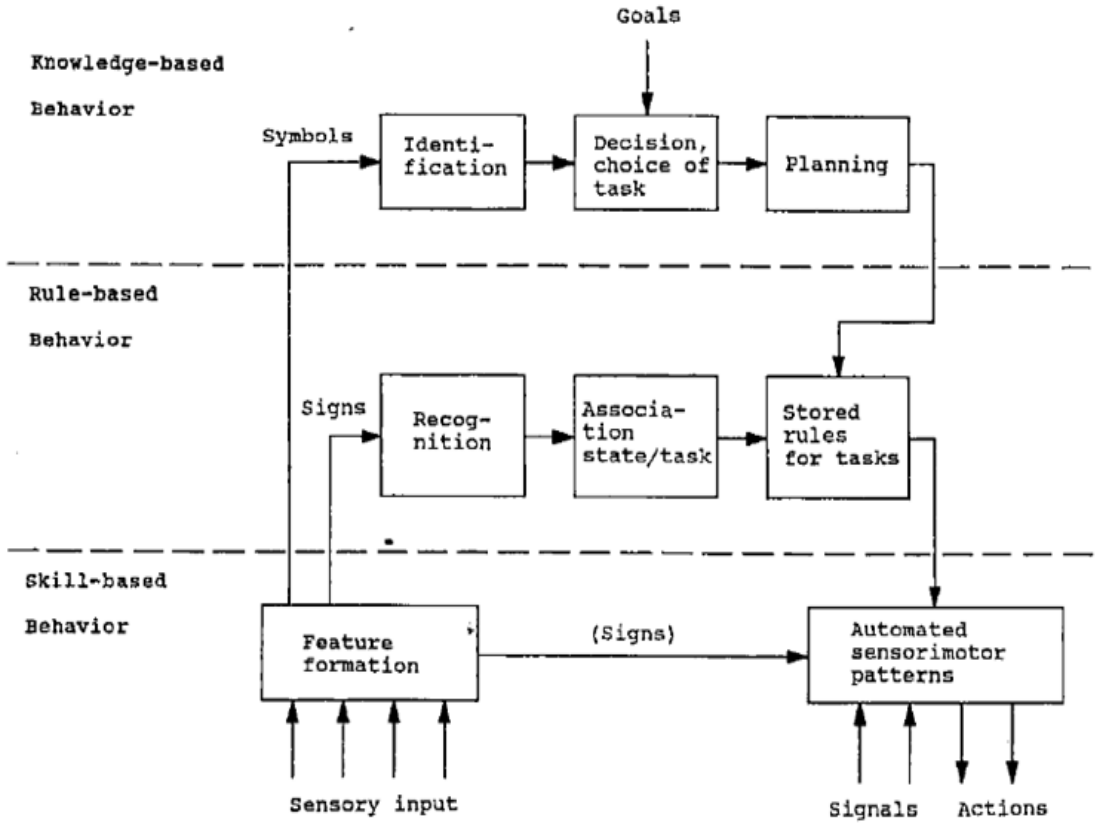
- Arbitration is a **subset** of **interaction** using the **input** and the **feedback** channel
- Arbitration consists of at least two phases: **moderation** and **decision**



- Arbitration requires the arbitration begin and the decision **criteria trigger**

Arbitration Levels

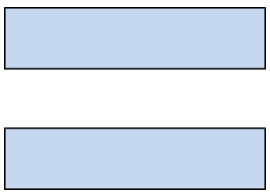


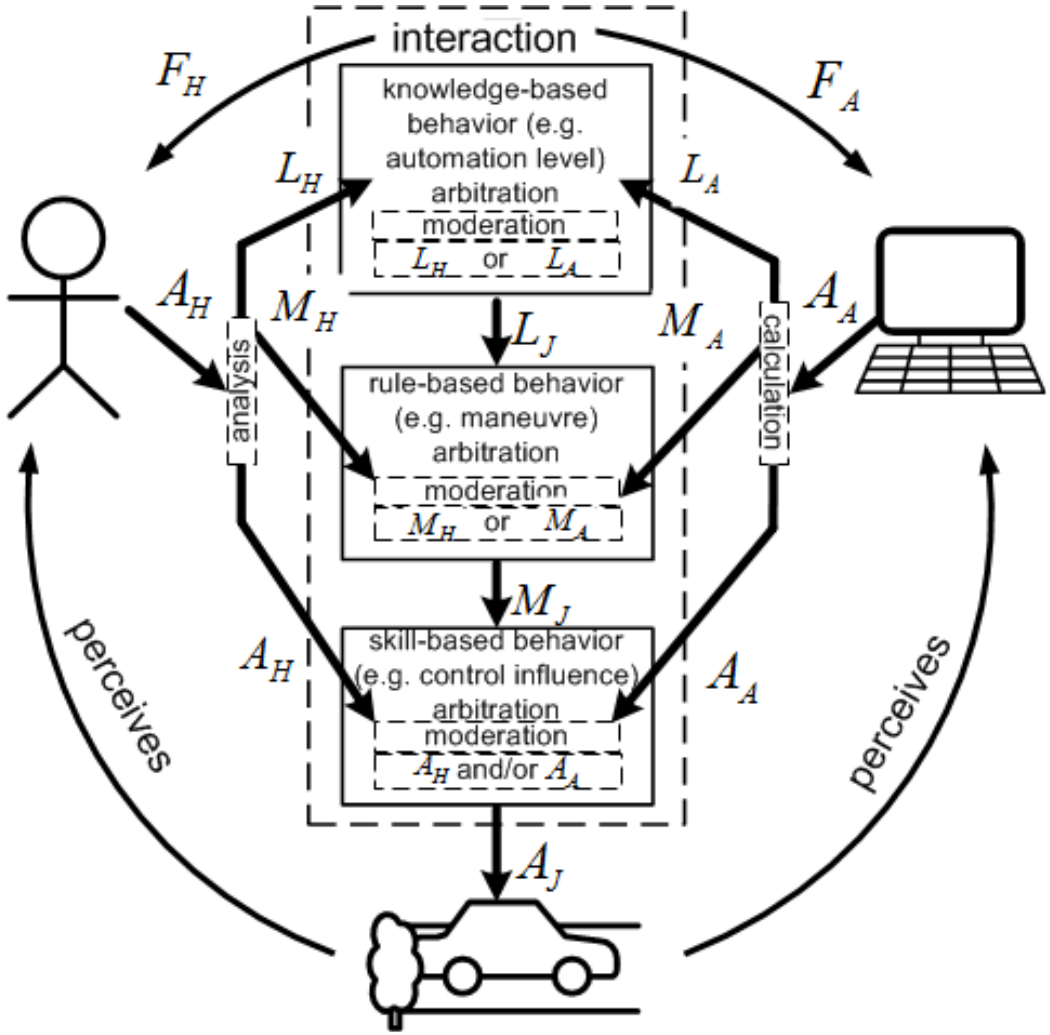


- Human **control behavior** model (*)
- **Knowledge**-based, **rule**-based, **skill**-based behavior

(*) after Rasmussen 1986

Arbitration Levels





interactive

Systematic notation and options for moderation

$$\begin{pmatrix} L_J \\ M_J \\ A_J \\ F_H \\ F_A \end{pmatrix} = f\left(\begin{pmatrix} L_H \\ M_H \\ A_H \end{pmatrix}, \begin{pmatrix} L_A \\ M_A \\ A_A \end{pmatrix}\right)$$

- Possible **notation** as mathematical (vector-)term (output = **function** of input)
- **Advantage**: very short, comprehensible, ready for possible calculations, transformations etc.). But **take care** – That‘s all **about design**!

$$\begin{pmatrix} L_J \\ M_J \\ A_J \\ F_H \\ F_A \end{pmatrix} = f\left(\begin{pmatrix} L_H \\ M_H \\ A_H \end{pmatrix}, \begin{pmatrix} L_{A1} \ldots L_{An} \\ M_{A1} \ldots L_{An} \\ A_{A1} \ldots A_{An} \end{pmatrix}\right)$$

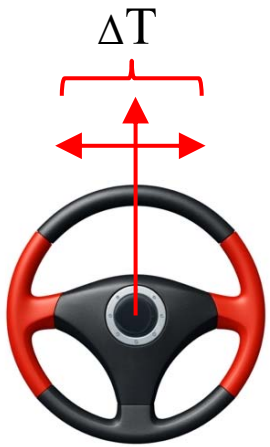
- For more successful moderation one require from automation **more then one option**
- For example: Automation is able to provide ‚manual‘ **as well as** ‚highly automated‘ mode and maneuvre ‚follow vehicle‘ is possible **as well as** maneuvre ‚change lane‘

Arbitration space perspective: Where to arbitrate?

- **Simple** way to arbitrate over a switch (e.g. turn indicator switches off the lane departure warning)
- Disadvantage: **more intuitive** sophisticated arbitration is **not possible**

- **Speech** recognition
- **Gesture** recognition
- **Facial expression** recognition
- **Eye-tracking**
- **Pressure, capacity** recognition
- Special arbitration **inceptors**

- Arbitration over the inceptor **angle/position**
- Disadvantage: without X-by-Wire direct **intervention in control**
- Solution: Use of (time triggered) **dead zone**



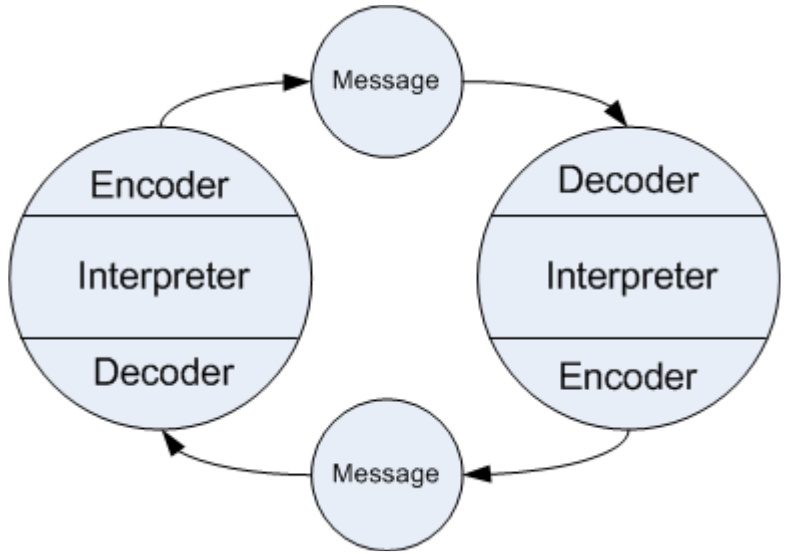
- Arbitration over the inceptor **torque/force**
- Disadvantage: **very noisy** under real conditions
- Solution: Use of (time triggered) **filtering** algorithm

What about interplay between interaction and behavior?

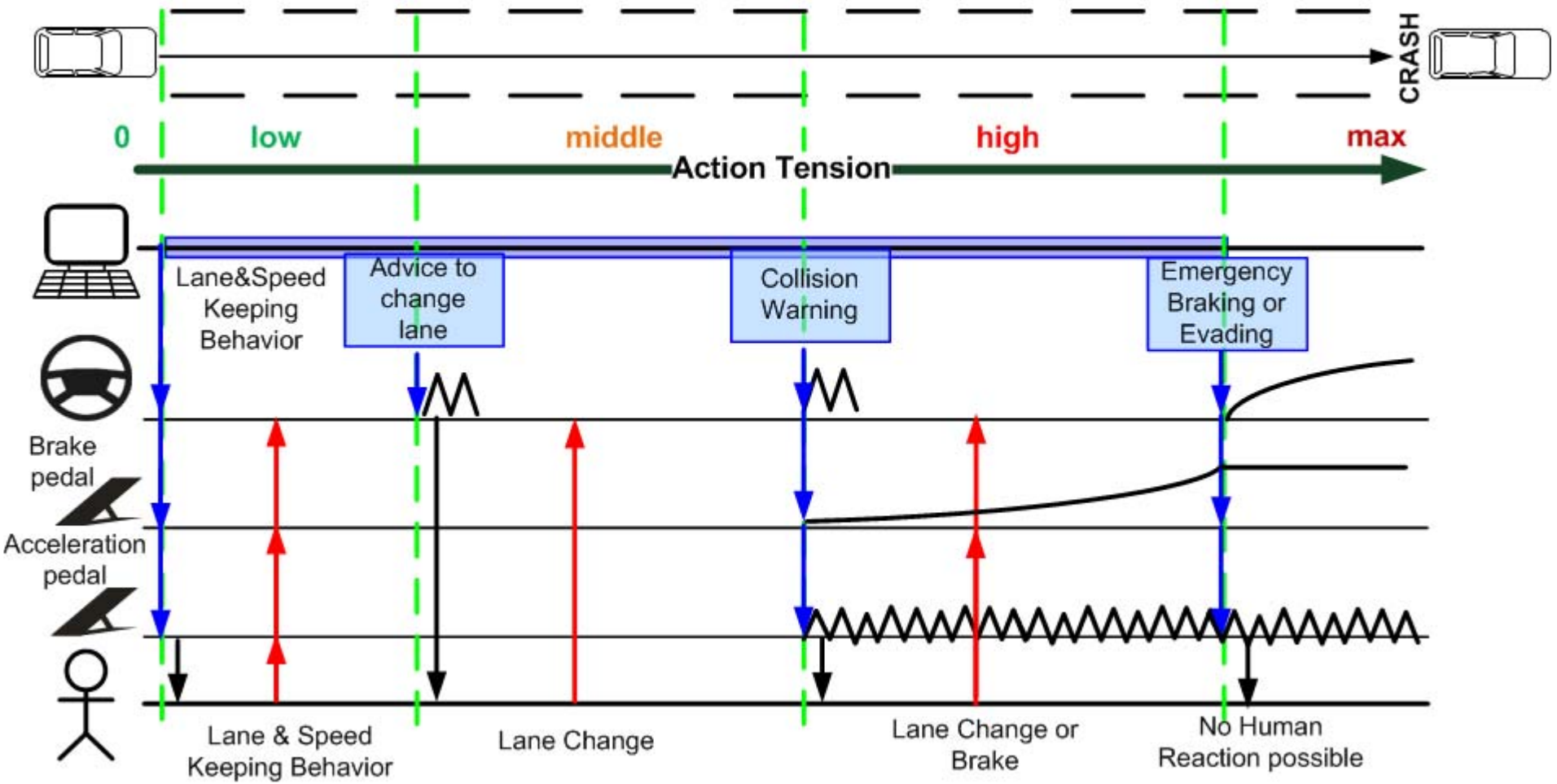
- Human and machine **behavior** causes **communication**
  - Axiom: One cannot not communicate (Watzlawick 1967)
- **Communication** is a nessesary precondition for the **interaction**
  - Communication/Interaction model by Shannon(1948)/Schramm(1954)
- **Interaction** (arbitration) influences the system **behavior**

=> **Match** of interaction and behavior design **is necessary**!

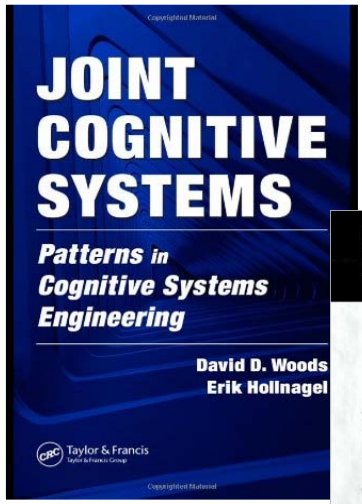
Picture source: wiehl.de

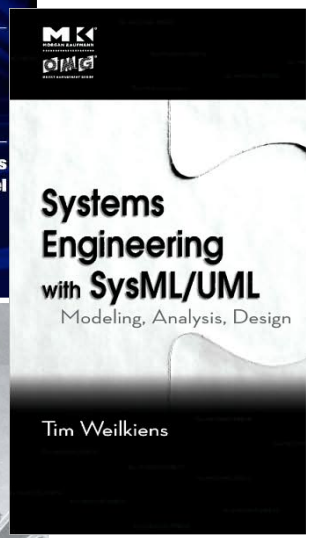


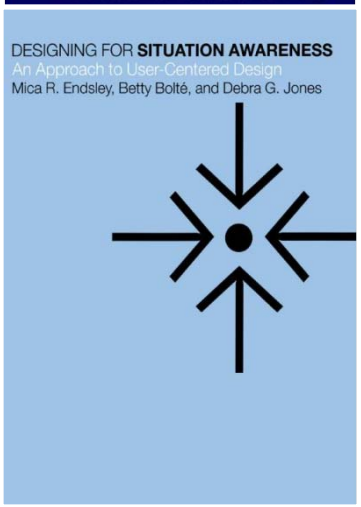
Arbitration design notation: sequence diagram

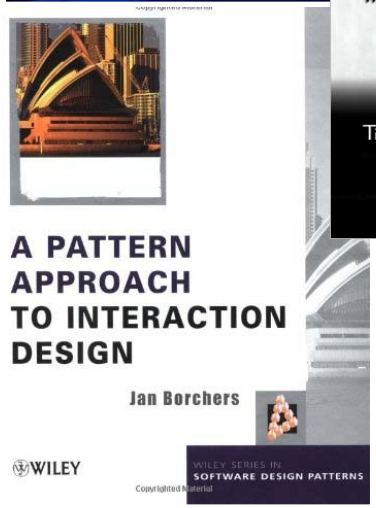


How a dynamic cognitive system can be designed using arbitration concept in general?



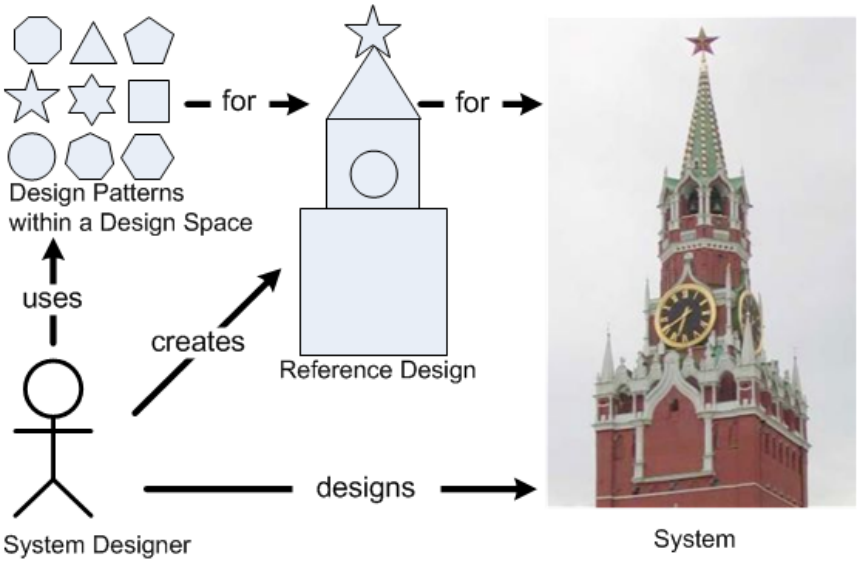






- Arbitration concept can be related to the agent behavior modeling part of **cognitive systems engineering**
- Systems Engineering **Methods and tools** (e.g. SysML (UML) e.g. sequence diagram, System-FMEA)
- Using arbitration strategies as **design patterns**

Design patterns



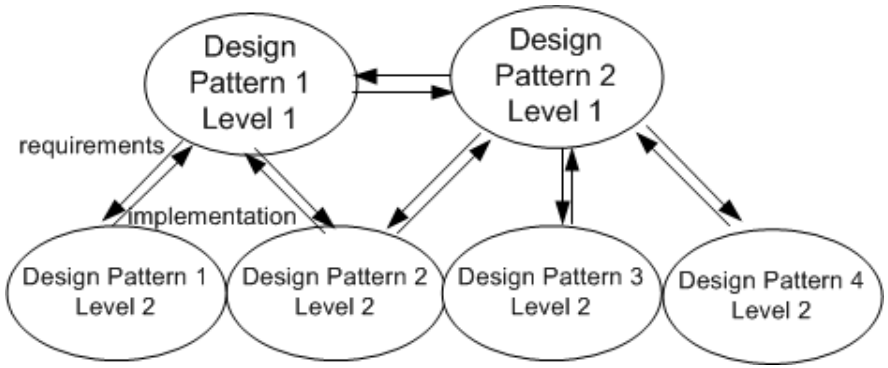
- A design pattern is a **proven solution for frequently occurring problems** in the design of architectures […] (Alexander 1977)
- In the last years the design pattern approach also got very common in the **area of human computer interaction**

ARTEMIS	**D3CoS** Designing Dynamic Distributed Co-operative Human-Machine Systems	D3CoS
D3-03 **Reference Designs and Design Patterns for Cooperation & DCoS State Inference and Adaptation** **Preliminary Version**		

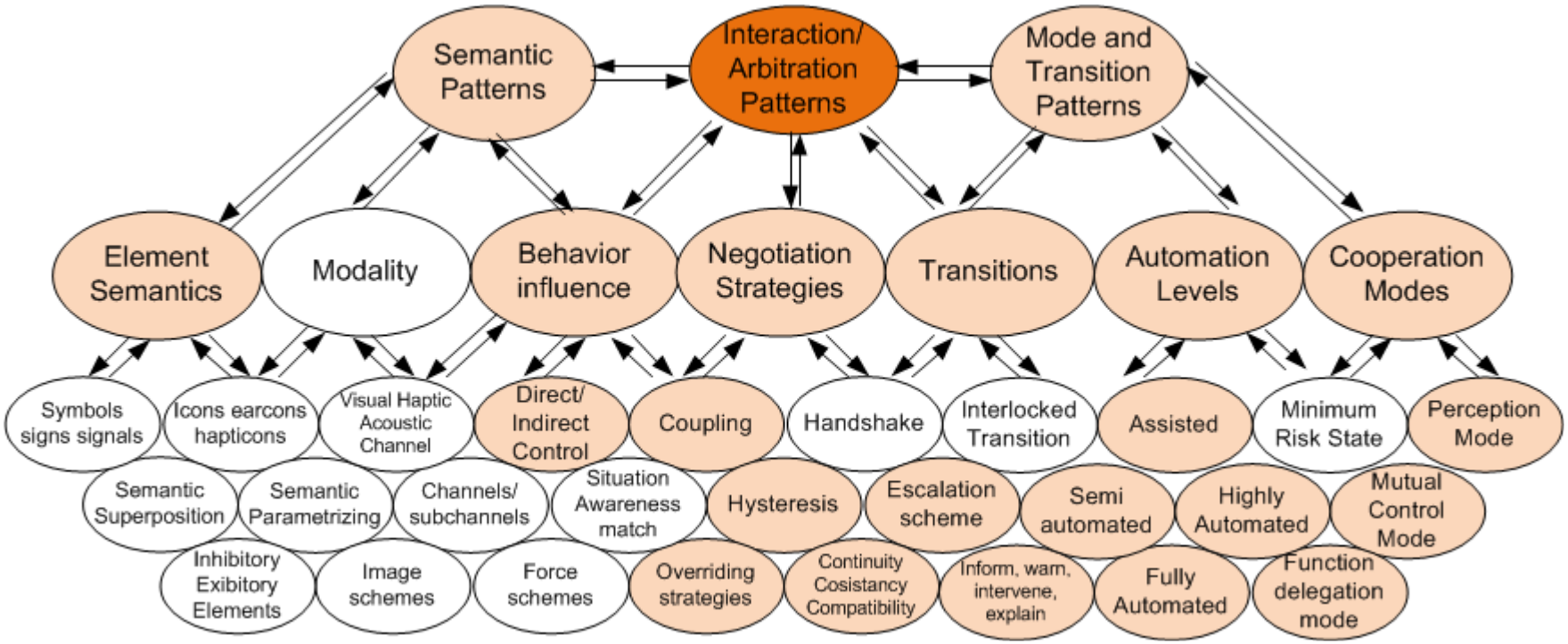
- Reference Designs and **Design Patterns for Cooperation** & DCoS State Inference and Adaptation (*)

(*) EU-Project D3CoS, public deliverable D3-03, 2012

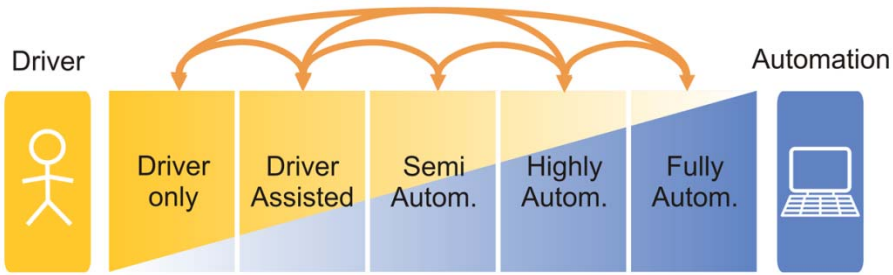
Design patterns as a framework for ‚arbitration pieces‘

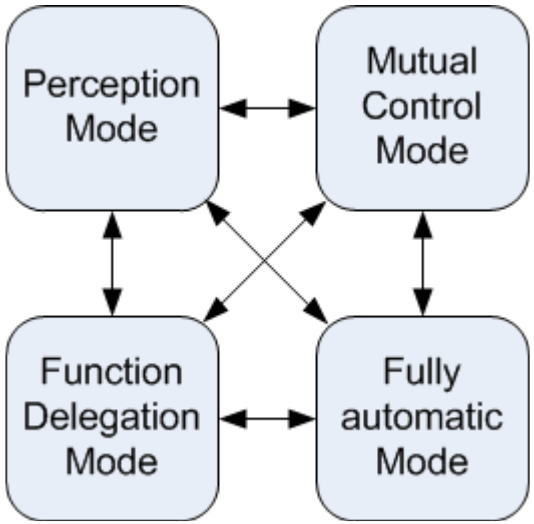


- Design patterns can be organized **hierarchicaly**
- Upper levels are more abstract, lower levels are more concret
- Upper levels **require** lower levels, lower levels **implement** upper levels



Design Patterns: System behavior states and transitions, Automation Level, Cooperation Modes and Transitions

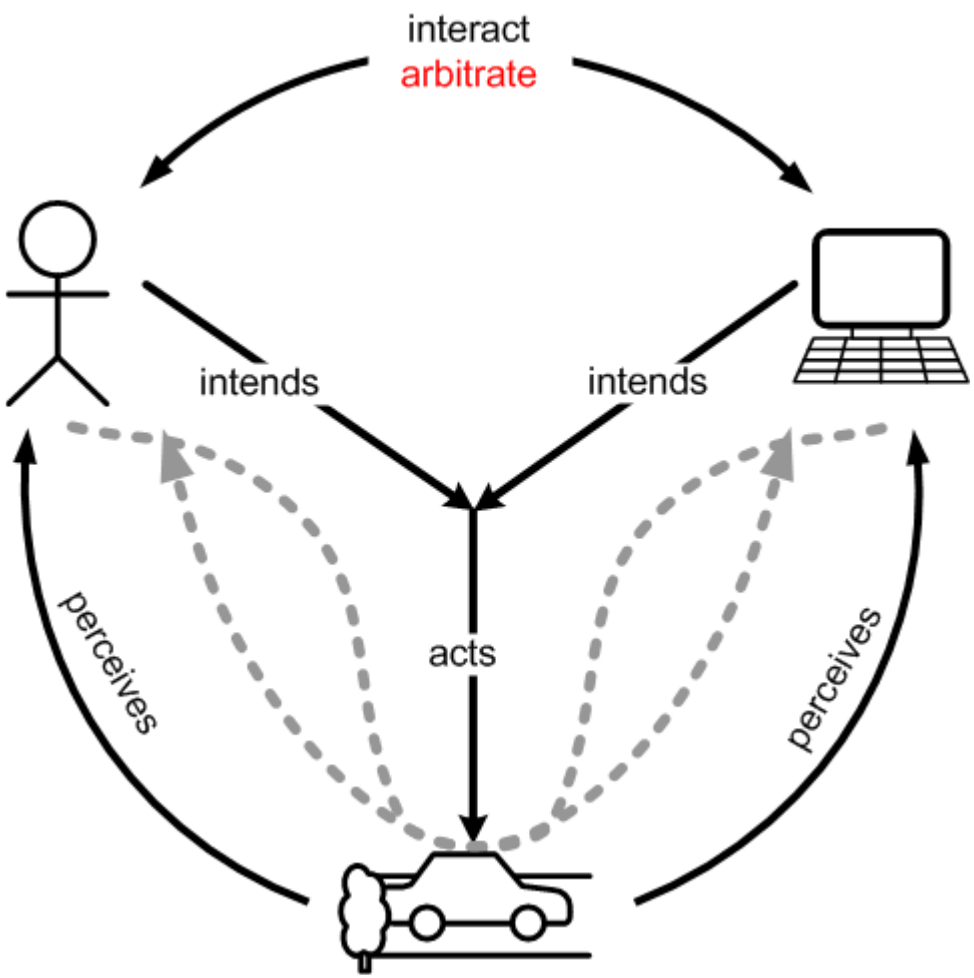




- **Automation level** (*,**) as more **quantitative** (how much automation is there?) **state-machine-like** description of the control distribution between human and automation with **transitions** (***) between states
- More **qualitative** point of view (who does what and how?) and task allocation oriented human-machine behavior description with '**cooperation modes**' (****)
- **Both perspectives** (quantitative and qualitative) are important and **compatible** to each other (*****)

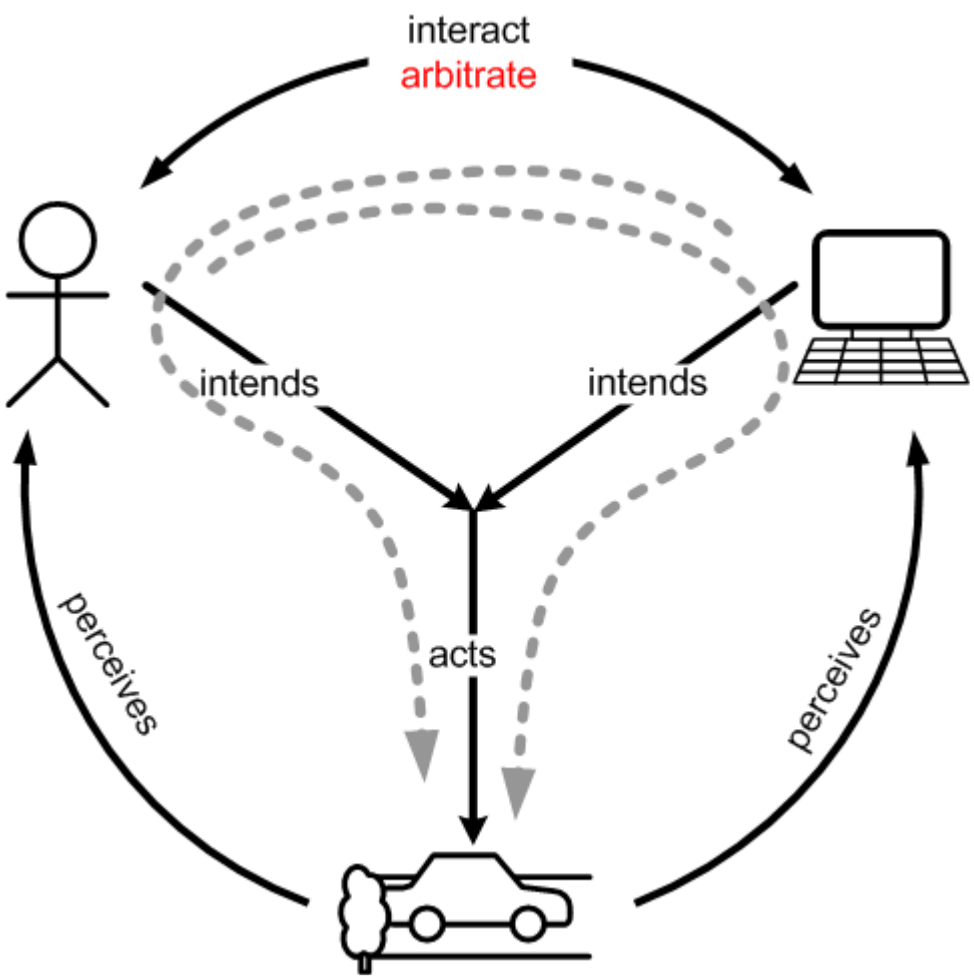
(*) Parasuraman et al. 2000 (**) Flemisch et al. 2008 (***) Schieben et al. 2011 (****) Hoc 2001 (*****) EU-Project D3CoS public deliverable D3-03

Direct behavior influence by implicite communication



- Human controls directly, automation gets an implicite feedback
- Automation communicates implicitly by direct control
- Human gets the feedback and can correct the intended machine action
- Used by **active elements**: For example active steering by lane keeping assist

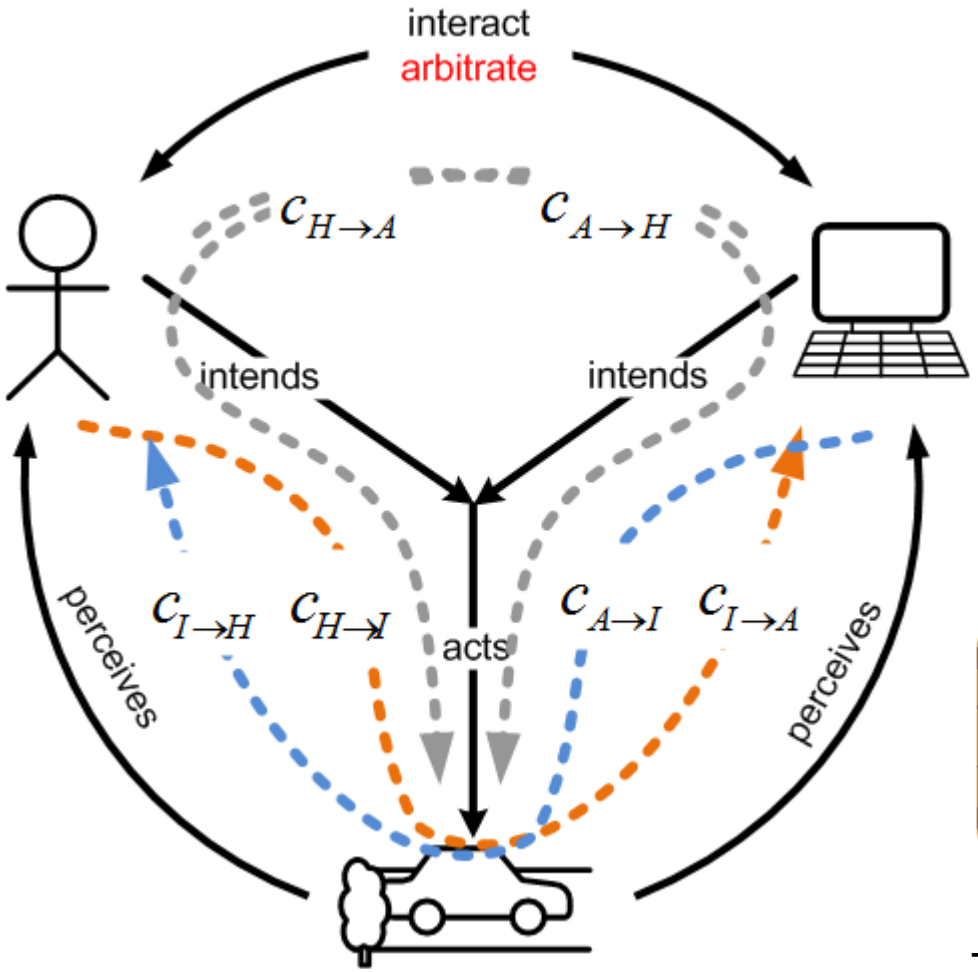
Indirect behavior influence by explicite communication



- Automation communicates **explicitely** by trying to control **indirectly**
- Human gets the information and can implement the intended machine action or not
- Used by **passive elements**: for example vibration by lane departure warning assist
- **Indirect** control by human is imaginable, for example conduct-by-wire concept (vehicle guiding **over maneuver triggering**) (*)

(*) Winner & Hakuli 2006

Behavior influence by coupling(*) (of direct/indirect control intent)



- **Control influence** and **Control feedback** one can describe with **coefficients**

$$c_{X\to Y} \in [0..1]$$

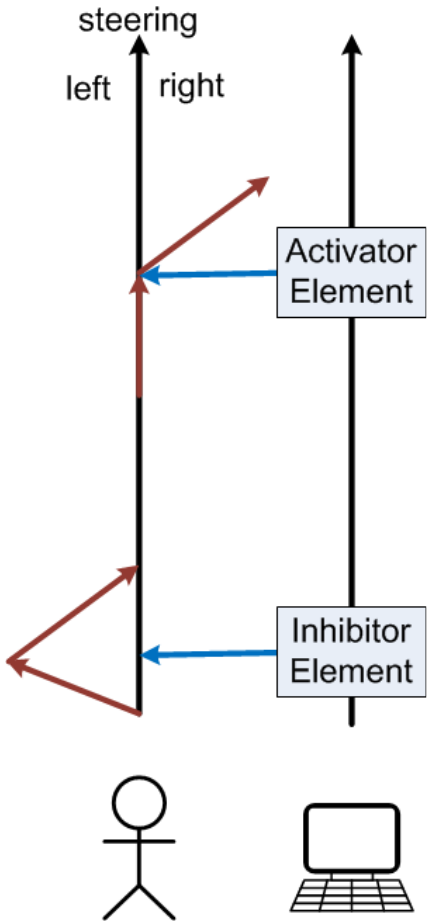
- **H → Human**
- **A → Automation**
- **I → Interface**

$$\begin{pmatrix} A_J \\ F_H \\ F_A \end{pmatrix} = \begin{pmatrix} A_H * c_{H\to I} + A_A * c_{A\to I} \\ A_A * c_{A\to I} * c_{I\to H} \\ A_H * c_{H\to I} * c_{I\to A} \end{pmatrix}$$

- **Indirect control** intent coupling is **possible**

(*) First ideas described in Flemisch et al. 2010

Behavior influence by element semantics (inhibitors/exhibitors)



- Let‘s imagine that the human is **not steering at all**
- There is an **interaction element** possible **activating** the human behavior
  - For example active steering intervention by a lane keeping assist that is aimed to **activate** human steering behavior

- Let‘s imagine that the human is **steering left**
- There is an **interaction element** possible **inhibiting** the human behavior
  - For example vibration on the steering wheel during leaving the lane by a lane departure warning assist that is aimed to **inhibit** the human steering behavior

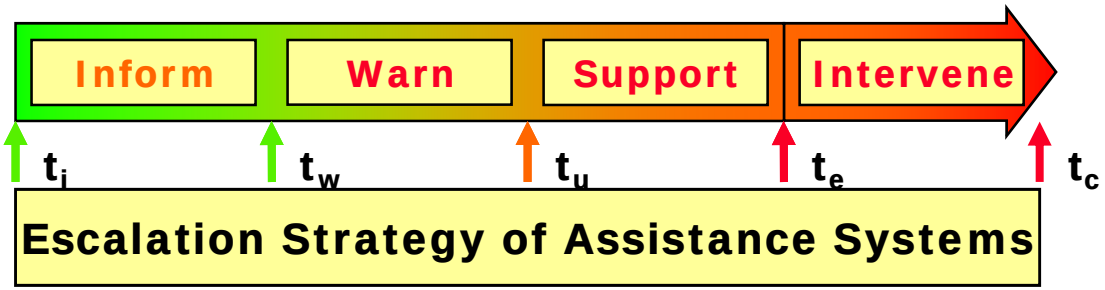
Behavior influence by basic element semantics (semantic toolbox)

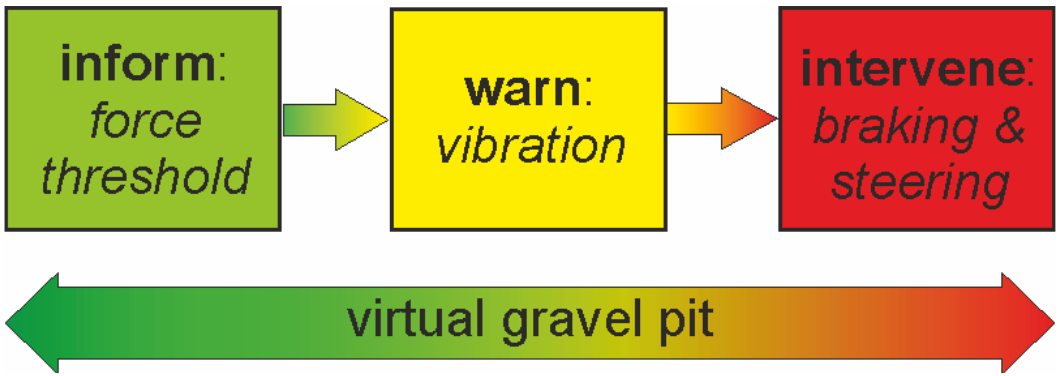
- **Consistency** in semantics is essential
- Interaction elements must be **unambiguous**

Modality	Element Name	Element Details	Meaning
haptic	Ticks	Directed pulses of low to medium amplitude	directed hint, suggestion, advice or request to act/decide
haptic	Vibration Pulse	vibration of short duration with changing amplitude or frequency	Indication of Status, Mode or Transition
haptic	Vibration	medium to high frequency, low amplitude	Indication of negative valency or danger
haptic	Shake	Low frequeny, medium to high amplitude, undirected	Imminent danger, triggers reflexive control take over
haptic	Continuous Force / Torque		Feedback of automation action, action execution or correction (away from negative valency or danger)
haptic	Blockage / Force barrier	force threshold (a certain force must be exceeded before inceptor moves beyond certain position)	directed action resistance (by automation)
haptic	Position Offset		Feedback of autmation action
haptic	Stiffness		Feedback of automation / assistance level
haptic	Kineasthetic Pulse	Short dynamic change of motion that can be felt withoutlarge impact on movement of vehicle?	emergency warning and suggestion of imminent action
acoustic			
visual			

interactive

Behavior influence by complex element semantics





- There is a well-known **escalation strategy** used in design of driver assistance systems
- Using that strategy and simple semantics we are able to **design more complex elements (*)**
- **Virtual gravel pit** as an element with the complex semantic feels and takes effect like a real gravel pit

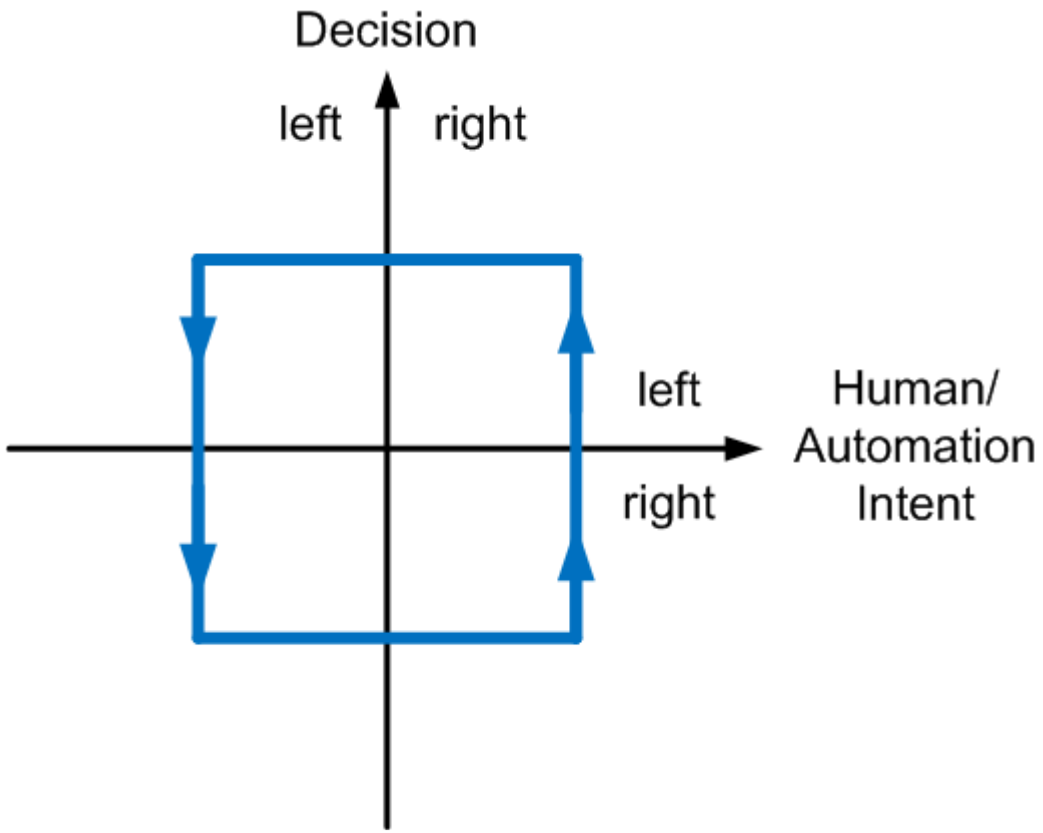
(*) First ideas already described in Kelsch et al. 2006

Generic arbitration strategies and rules

Image sources: kooperation-mediation.de, 2012, blog.team-vision.at 2012

- There are some **generic** arbitration strategies which can be derived from well know **concepts or metaphors**, for example:
  - **Dialog** rules like let finish the speaking, be polite, pay attention etc.
  - **Clarify** the **roles** and and the **authority** distribution
  - Support behavior **reliability** and **trust**
  - Arbitrate **consistently** (continuous), **compatibely** (inner- and outer compatibility), **transparently** (comprehensible, observable)
  - **Etc.**

Example of a specific arbitration strategy: Hysteresis



- Frame of reference with the **dependancy** between the human/automation **intent** and arbitration **decision**
- Simulation of the **inertial behavior** with the help of **hysteresis (*)**

(*) First ideas already described in Löper et al. 2008

Development Tool: Arbitration – Overriding Matrix (*)

Calculated Automation intent Analyzed Human intent	L_A = Automation Level Y M_A = Maneuver Y A_A = Action Y
A_H = Action X M_H = Maneuver X L_H = Automation Level X	**system behaviour:** continues or discontinues L_J = Automation Level Z M_J = Maneuver Z A_J = Action Z **interaction:** interaction description in words F_H = A, F_A = B

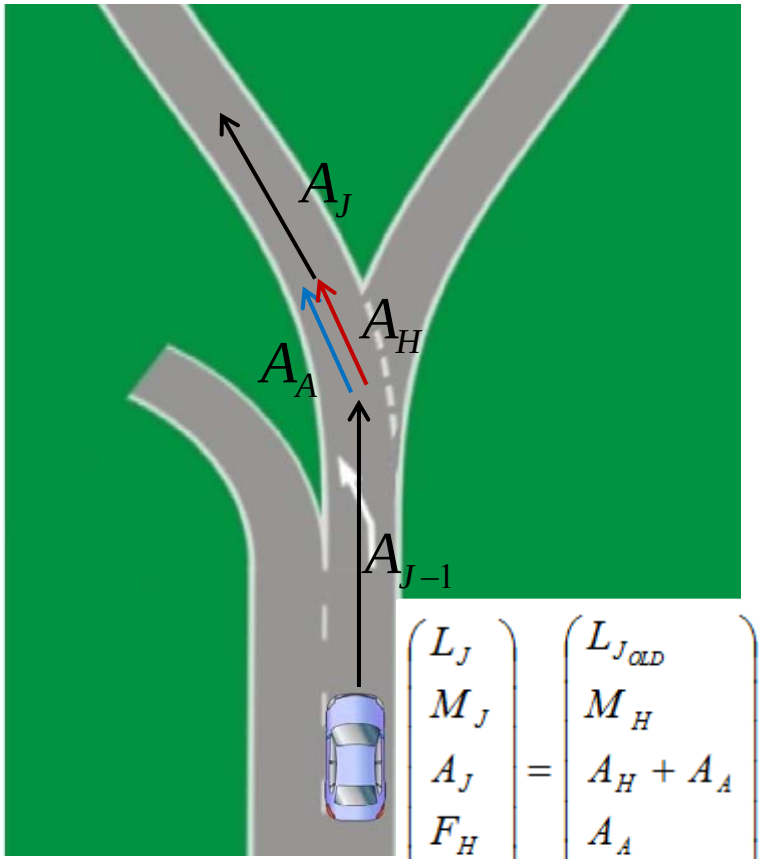
- **You remember** the mathematical expression?

$$\begin{pmatrix} L_J \\ M_J \\ A_J \\ F_H \\ F_A \end{pmatrix} = f\left(\begin{pmatrix} L_H \\ M_H \\ A_H \end{pmatrix}, \begin{pmatrix} L_A \\ M_A \\ A_A \end{pmatrix}\right)$$

- For different use-cases a **matrix-like notation** is possible
- The cell contains the notation for the **design function**

(*) InteractIVe public deliverable D3.2 / IWI Strategies

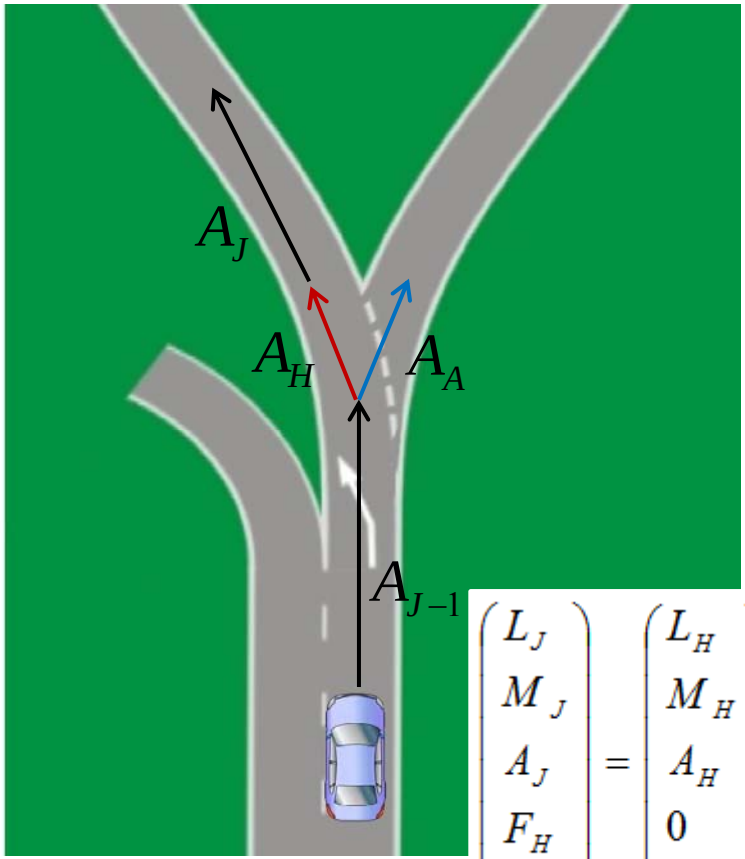
Arbitration – Overriding Matrix: No Conflict



$$\begin{pmatrix} L_J \\ M_J \\ A_J \\ F_H \\ F_A \end{pmatrix} = \begin{pmatrix} L_{J_{OLD}} \\ M_H \\ A_H + A_A \\ A_A \\ A_H \end{pmatrix}$$

Calculated Automation intent Analyzed Human intent	L_A = keep actual automation level M_A = turning left A_A = steering left
A_H = steering left M_H = turning left L_H = keep actual automation level	**system behaviour:** continues (driver agree) $L_J = L_{J_{OLD}}$ $M_J = M_H = M_A$ = turning left $A_J = A_H + A_A$ **interaction:** feedback about actual automation level and maneuver, continue steering torque on active steering wheel $F_H = A_A$, $F_A = A_H$

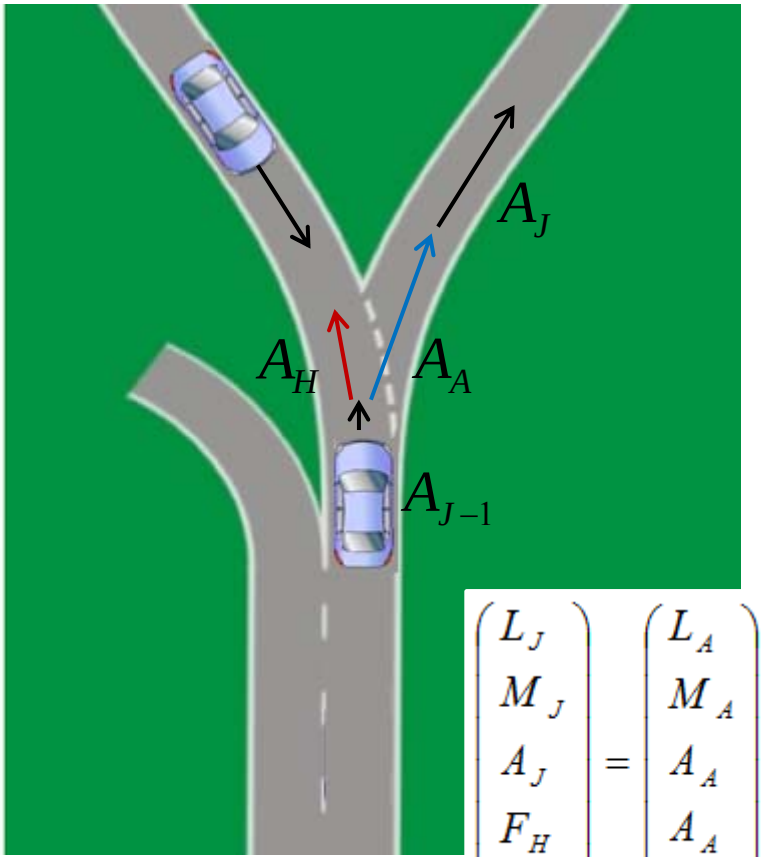
Arbitration – Overriding Matrix: Conflict



$$\begin{pmatrix} L_J \\ M_J \\ A_J \\ F_H \\ F_A \end{pmatrix} = \begin{pmatrix} L_H \\ M_H \\ A_H \\ 0 \\ A_H \end{pmatrix}$$

Calculated Automation intent Analysed Human intent	L_A = keep actual automation level M_A = turning right A_A = steering right
A_H = steering left M_H = turning left L_H = manual control	**system behaviour:** discontinues (driver disagree) $L_J = L_H$ = manual control $M_J = M_H$ = turning left $A_J = A_H$ **interaction:** feedback about new choosen automation level and maneuver, no steering torque on the steering wheel $F_H = 0$, $F_A = A_H$

Arbitration – Overriding Matrix: Strong Conflict



$$\begin{pmatrix} L_J \\ M_J \\ A_J \\ F_H \\ F_A \end{pmatrix} = \begin{pmatrix} L_A \\ M_A \\ A_A \\ A_A \\ A_H \end{pmatrix}$$

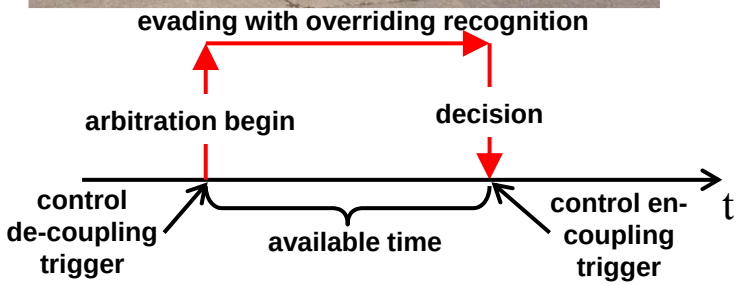
Calculated Automation intent Analysed Human intent	L_A = fully automated M_A = emergency evading right A_A = strong steering right
A_H = steering left M_H = turning left L_H = keep actual automation level	**system behaviour:** discontinues (automation disagree) $L_J = L_A$ = fully automated $M_J = M_A$ = emergency evading right $A_J = A_A + (A_H = 0)$ **interaction:** feedback about new choosen automation level and maneuver, continue steering torque on active steering wheel $F_H = A_A$, $F_A = A_H$

How to find the best interaction (arbitration) design?

- Explorations and experiments are the **baseline** for the ‚good' interaction (and arbitration) design
- In **InteractIVe** some interaction designs for automatic **steering intervention in emergency** situation were tested in simulator experiments (*)
  - Some independent variables were tested: **Intervention torque** (4.5-9.9 Nm), several **preliminary warning strategies**
- Interaction (arbitration) **design implications**:
  - **Trade-off** between overridability and successful automation behavior was not found yet
  - **Driver behavior** during the evading **interferes negatively** with automation behavior (lateral as well as longitudinal)
  - Adequate **trigger for begin of arbitration** under simulation condition was not found yet

(*) InteractIVe public deliverable D3.2 / IWI Strategies

How to find the best interaction (arbitration) design?



- **Now** in DLR3 we are setting up an experiment in a same scenario under more real conditions in a test vehicle (FASCar II)
- Still automatic **steering intervention in emergency** situation, but:
  - **Former expert study** to get an (local) optimum in interaction design
  - Arbitration design using **temporal driver de-coupling** by steer-by-wire **with overriding possibility vs. no driver decoupling** → two different arbitration designs
  - **Systematic testing** of automation **plus** human performance by **correct vs. false** triggering of the evade maneuver

interactive

Outlook

- **Empiric studies** for evaluation of arbitration patterns are necessary
- Development of arbitration tools towards a compilation of an arbitration **toolchain**
- Further **theoretical work**
  - Development of the **consistent mathematical and graphical description** of arbitration patterns
  - Further development of **supplementary concepts** like action-tension concept (*)
  - Development of the arbitration concept towards a human-machine **interference** concept

(*) Will be published on EAM 2012 in Kelsch et al. 2012, Sept.

References

EU-Project D3CoS: Reference Designs and Design Patterns for Cooperation & DCoS State Inference and Adaptation, Public Deliverable D3-03, 2012

EU-Project InteractIVe: IWI Strategies, Public Deliverable D3.2, 2012

Flemisch, F., Heesen, M., Kelsch, J., Schindler, J., Preusche, C., Dittrich, J. (2010): Shared and cooperative movement control of intelligent technical systems: Sketch of the design space of haptic-multimodal coupling between operator, co-automation, base system and environment. 11th IFAC Symposium; Valenciennes, France, 2010

Flemisch, F., Kelsch, J., Löper, C., Schieben, A., Schindler, J. (2008): Automation spectrum, inner / outer compatibility and other potentially useful human factors concepts for assistance and automation, Ed. de Wart; Flemisch: Human Factors for Assistance and Automation; Shaker Publishing, 2008

Griesche, S., Kelsch, J., Heesen, M., Martirosjan, A. (2012): Adaptive Automation als ein Mittel der Arbitrierung zwischen Fahrer und Fahrzeugautomation. AAET 2012, Braunschweig

Hoc, J. M. (2001). Towards a cognitive approach to human-machine cooperation in dynamic situations. International Journal of Human-Computer Studies, 54, pp. 509-540

Kelsch, J.; Flemisch, F.; Löper, C., Schieben A., & Schindler, J. (2006). Links oder rechts, schneller oder langsamer? Grundlegende Fragestel-lungen beim Cognitive Systems Engineering von hochautomatisierter Fahrzeugführung. In: 2006 Cognitive Systems Engineering in der Fahrzeug- und Prozessführung DGLR-Bericht 2006-2. S. 227-240

Löper, C., Kelsch, J., Flemisch, F. (2008): Kooperative, manöverbasierte Automation und Arbitrierung als Bausteine für hochautomatisiertes Fahren. Gesamtzentrum für Verkehr Braunschweig (Hrsg.): Automatisierungs-, Assistenzsysteme und eingebettete Systeme für Transportmittel. GZVB, Braunschweig. pp. 215-237

Parasuraman, R., Sheridan, T. B., Wickens, C.D. (2000): A Model for Types and Levels of Human Interaction with Automation. IEEE Transaction on Systems, Man, and Cybernetics 30(3): pp. 286-297

Rasmussen, J. (1986): Information processing and human-machine interaction: An approach to cognitive engineering. p. 101, North-Holland, New York, ISBN: 0444009876

Schieben, A., Temme, G., Köster, F., Flemisch, F. (2011): How to interact with a highly automated vehicle—generic interaction design schemes and test results of a usability assessment. Human Centred Automation. D. de Waard, N. Gerard, L. Onnasch, R. Wiczorek and D. Manzey. Maastricht, Shaker Publishing: pp. 251-266

Schramm, W. (1954): How Communication Works. In: W. Schramm (ed.). The Process and Effects of Mass Communication. Urbana: University of Illinois Press 1955, 3-26

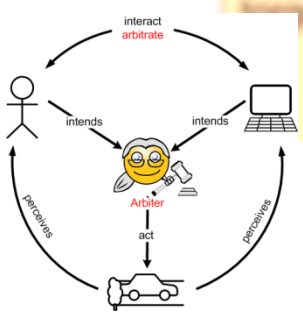
Shannon, C.E. (1948): A Mathematical Theory of Communication, Bell System Technical Journal, Vol. 27, p. 381

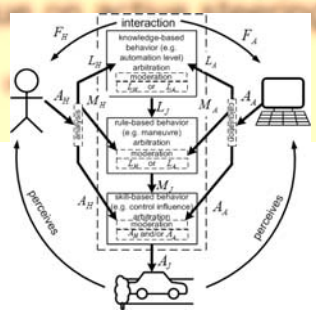
Watzlawick, P., Helmick Beavin, J., Jackson, D.D. (1967): Pragmatics of Human Communication: A Study of Interactional Patterns, Pathologies, and Paradoxes, Norton, ISBN: 0393010090

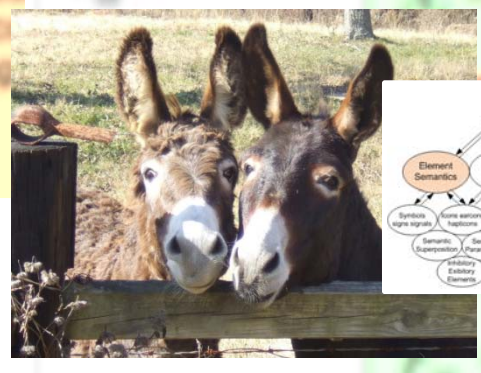
Thank You

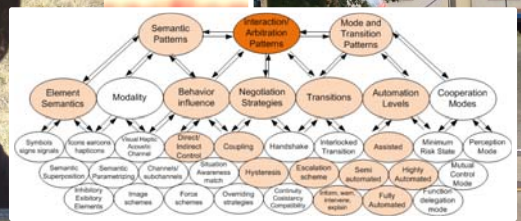
johann.kelsch@dlr.de

Let's arbitrate!









Picture Source: Reed A. Cartwright, 2008, pandasthumb.org