Active controllability control in worlds with discontinuous controllability switches

Marc Toussaint

May 13, 2011

Abstract. These are some informal and at some points preliminary notes on a research idea. I publish them as a reference for the 2011 BSc thesis "Learning Symbols for Hierarchical Control from Interaction: Controllability Control" by Gregor Gebhardt, which took these notes as the starting point.

1 Informal description of the structure we're after

In this work we follow the assumption that a certain, very special kind of structure is inherent in natural worlds. We will first describe this structure informally, then define it in mathematical terms, then define a toy benchmark world. Clearly, this kind of structure is not the only one in natural worlds – there are diverse kinds of structures in physical worlds – we only try to capture one aspect which, hopefully, will be helpful to organize goal-directed manipulation in natural environments composed of objects.

Consider a robot in a household environment. There are degrees of freedom the robot can directly actuate: its arms and hand, perhaps it position. And there are other degrees of freedom which it can not directly actuate: the position of objects, the position of doors and drawers and other kinematic degrees of freedom. But these latter DoFs can in principle be manipulated if the robot manages to "attach" to them:

- grasp an object – and its position becomes manipulable
- grasp a handle of a drawer – and it can be pulled or closed
- grasp the handle of a door – and it can be lowered
- lower the handle of a door – and the door itself becomes manipulable
- touch a light switch – and it becomes pushable
- turn the key to a drawer – and it becomes openable

All these are examples where a robot can, by active change of state, acquire manipulability of DoFs of the environment that would otherwise not be manipulable.

The transition between whether a DoF is manipulable or not is approximately discrete: either you touch an object or not; either you touch a light switch or not, etc.

I think this this is a very special and important structure of manipulation environments which has potential to be exploited to bootstrap *symbolic* representations of such environments: The discrete transitions between manipulability of DoFs are important symbolic events. Fulfilling complex manipulation tasks in principle requires to generate these events in a certain sequential order. Therefore these events might provide a suitable symbolic abstraction for behavioral planning.

These symbols are inherently grounded in a basic control theoretical sense. They can unambiguously be defined – e.g. as proposed in the next section. The hope is that taking these kinds of symbols as first level of grounded symbolic representations alleviates the typical grounding problem in robotics and the arbitrariness of symbol definitions in natural environments.

The *disc world* that we will define below is an attempt to define a minimalistic toy world which highlights this specific kind of structure and therefore may be used in first attempts to test methods and concepts. Behind all this stands the hypothesis and hope that this specific structure indeed also captures a fundamental aspect of natural worlds – and theory and methods developed in the toy world will transfer to natural worlds. If this will not be the case, I consider the idea of this assumed structure as well as the definition of the disc world as a failure.

2 A formal definition of controllability structure

We assume the world has many continuous degrees of freedom. Later we will use relational representations to capture these DoFs. But for simplicity, let us start using a direct propositional representation and assume the world has n DoFs $x \in \mathbb{R}^n$.

We assume the agent has freedom to set controls $u \in \mathbb{R}^d$. As usual, the world can be described by the system equation

$$\dot x = f(x, u) \tag{1}$$

or a stochastic variant thereof – we neglect stochasticity for now.

Let $\dot x \approx A(x_0)x + B(x_0)u$ be a *local* linearization of the system dynamics around the reference state x_0. In general, not all DoFs of x are controllable. In/around state x, the matrices $A(x)$ and $B(x)$ define a classical notion of controllability: The system is controllable if the matrix

$$C = [B \; AB \; A^2B \; \cdots \; A^{n\text{-}1}B] \tag{2}$$

has full rank n. C is a $n \times dn$-matrix. Let

$$C = \begin{pmatrix} c_1^\top \\ \vdots \\ c_n^\top \end{pmatrix} \tag{3}$$

where c_i is the ith row of C. Intuitively, the row c_i corresponds to "how the ith DoF can be controlled" either directly by u via B or indirectly via AB, A^2B, etc. If two rows are linearly dependent, then the two respective DoFs can only be controlled dependently, not independently. Let $\hat{C} = CC^\top$ be the $n \times n$-matrix with $(CC^\top)_{ij} = c_i^\top c_j$. The matrix C as well as the symmetric $\hat{C}$ has full rank n if all c_i are independent. Computing the SVD $\hat{C} = USV^\top$, the rows of V with non-zero singular value span the space of controllable DoFs.

NOTE: We could continue here with defining the controllability state as the set of basis vectors $\{v_1, .., v_m\}$ that spans the current space of controllable DoFs. However, these vectors may continuously rotate depending on the state x (as is the case in the disc world) — we would loose the notion of discontinuous switches between distinct controllability states. My hypothesis remains that in natural worlds, although the basis vectors v_i might rotate, there is still certain dimensions (aligned with the original coordinate system) totally switched off and others not: they correspond to different "objects". I don't know yet how to formalize this properly... to be continued...

Let us define the *controllability state* $c(x) \subseteq \{1, .., n\}$ (short c-state) as the set of DoFs of that are controllable (according to $A(x)$ and $B(x)$) in state x. Let $n(x) = |c(x)|$ be the number of controllable dimensions in state x and $x_{c(x)} \in \mathbb{R}^{n(t)}$ the controllable state.

The definition of controllability is a discrete property: either a DoF is controllable or not. Therefore, the transitions between different c-states will be *discontinuous*.

Let

$$C = \{c(x) \ : \ x \in X\} \tag{4}$$

be the set of all possible c-states. This is a discrete set. We also assume it to be finite. (I think it would have to be some kind of fractal world for this to be infinite while X is bounded.) We can define a connectivity structure on this set as follows: Define

$$[c] = \{x \ : \ c(x) = c\} \tag{5}$$

as the set (equivalence class) of all states x that have the same c-state. These sets are disjoint. However, for convenience, let us make a notational shortcut and understand $[c] = [c] \cup \partial[c]$ as their closed sets (the sets including their closure). Then we say, two c-states c and d are adjacent iff

$$\gamma(c, d) = 1 \iff [c] \cap [d] \neq \emptyset \tag{6}$$

where γ is the adjacency function.

We call the graph structure (C, γ), which gives the set of all c-states and their adjacencies, the c-graph. It tells us whether we can in principle transition from a c-state c to another c-state d.

NOTE: This is not quite true yet – to be really able to actively transition we need to be able to control those DoFs of x that "penetrate" the separating manifold between $[c]$ and $[d]$.

Finally, for any neighboring c-states c and d let us define the separating set as

$$[c, d] := [c] \cap [d] \ . \tag{7}$$

These sets (which are not necessarily manifolds) are very interesting: they describe the set of states on which a discontinuous transition between c-states c and d is happening.

Problem definition: For each c-state learn the system dynamics within $[c]$. For each separating set $[c, d]$ learn how to transition it from $[c]$ to $[d]$ and vice versa. Learn to control on the symbolic level of the c-graph itself, in the sense of hybrid control.

3 The *disc world* as basic example

To exemplify the structure defined above we sketch a simple *disc world*.

Consider a 2D environment. There is an agent at position $y_0 \in \mathbb{R}^2$ and n objects at positions $y_i \in \mathbb{R}^2$, $i = 1, .., n$. Each object/agent is defined by its radius r_i, its mass m_i, and its friction coefficient f_i with $i = 0, .., n$ (the agent has index 0).

The agent can directly apply forces $u = m_0 \ddot{y}_0$ on its ball.

Contacts are handled as follows: if two balls collide, first the exact point in time of contact and the corresponding positions are computed. Then instantaneous impulse exchanges are computed with coefficient a for the elastic and b for the inelastic impulse exchange. Then it is decided whether the two balls remain in contact (HOW?). If yes, they will be simulated as kinematic chain, if no, they will be simulated as separated balls. Finally, the remainder of the simulation time step is integrated.

Variants:

- The world may be kinematic instead of dynamic (the state is $x = (u, y_{1:n})$ and controls are $u = \dot{q}$);
- The initial positions of objects may be randomized.
- There may be obstacles in the environment.
- The sizes/masses/frictions of objects/agent may differ and be randomized.