What are goals? And if so, how many?

Matthias Rolf and Minoru Asada
Osaka University, Japan
{matthias,asada}@ams.eng.osaka-u.ac.jp

***Abstract*—Goals are concepts used in many different areas of robotics, artificial intelligence, psychology, neuroscience, and also philosophy. Despite the wide usage, there is no common definition of a "goal". Rather, the term is used in substantially different ways even within disciplines. This paper discusses these notions and potentially unified views on goals, and points out how different perspectives on the same term lead to different arguments and can cause communication difficulties in the interdisciplinary community. We discuss how far goal terminologies can be generally considered as desired end states of action and point out the pivotal aspect of their explicit representation. As a major point we discuss the relation of such goals with reward and value systems from various perspectives.**

***Index Terms*—Goals, Goal Systems Development, Reward and Value Systems, Motivation**

I. Motivation

The concept of goals is used in many different disciplines to describe, model, or analyze the behavior and function of natural or artificial agents. Goals are used to describe simple animal behavior towards objects [1]. In psychology goals are considered to steer action [2] and facilitate learning [3], but also provide a mechanism to understand others' action [4] – so in neuroscience [5], [6]. In philosophy goals are part of a large discussion about intentionality and free will [7]. Research on artificial intelligence involves all of these aspects. Consequently, goals have long played a crucial role in AI systems and architectures [8].

In Developmental Robotics and Autonomous Mental Development goals have long played a less significant role – driven by the criticism of traditional AI agents with handcrafted goals and specific purposes chosen by a designer [9]. In fact, early psychologists like Piaget [10] denied a possible role of goals in very early development. This view has been revised [11], though, due to indications of early goal-directed behavior even in neonates [12], [13]. Therefore, both the role of goals *for* development as well as the development *of* the goals themselves have become research targets in developmental robotics [14], [15], [16]. The overarching, interdisciplinary character of artificial intelligence and developmental robotics requires to deal with conceptualizations from all of the other disciplines. Such understanding is very difficult in the case of goals. Goals come with a vast variety of (mostly implicit) definitions even within disciplines. In different disciplines the term is often used with very different meanings, premises and implications, so misinterpretations are common.

The primary aim of this article is to point out these differences, raise awareness of them, and then work towards a common understanding of these facets. We focus particularly on notions of goals and their relation to rewards and values, and discuss how far they can be regarded as distinct, or different perspectives on the same thing, or even synonymous. In Sec. II we point out the core meaning of goals as desired end states, which conforms with most, but not all, usages in natural speech as well as research literature. For deviating cases we discuss whether they can and should be still interpreted as an instance of an overarching "goal" concept. In Sec. III we discuss notions of such end-states that come with largely different implications, such as an agent being intentional in a philosophic sense, or indicating the mere presence of a physical object. We analyze these notions and suggest that a large variety of them can be explained by different perspectives on the same matter. We point out that if goals and rewards are kept as distinct concepts, they preserve an intimate relation that can be well analyzed by means of these different perspectives. We extend this thought in Sec. IV and discuss a possible emergence of goal representations out of reward signals as pointed out in previous research, and analyze assumptions and consequences of regarding such relation as right or wrong. In Sec. V we then discuss what such goal representations might do more, better, or different for an agent. Finally, we give concluding remarks in Sec. VI.

II. Goals: Desired End-States of Action?

What is the most common, the core meaning of a "goal"? In order to approach a scientific definition it is worth also considering the general use of such a word. Dictionaries – across languages[1] – refer to the word as some

1) "state of affairs", a "result", or an "object"
2) concerning someone's "action", "ambition", or "plan"
3) to which that ambition is "directed", or which is "desired" or "intended" to be reached.

The first point can be generally interpreted as some *state*, or set of states of the world, in which certain criteria are true. The second point – throughout languages – points out the immediate relation to an action. We have previously pointed out [16] that this distinguishes goals from wishes, which are also desired states, but at least not necessarily related to concrete action or planning. The third point states the direction of the action: towards that state. It is worth mentioning that *none* of these definitions mentions the *reason* for the underlying intention, which is different from "purposes".

[1] oxforddictionaries.com "goal"; corresponding definitions in other languages: duden.de (German) "Ziel"; nlpwww.nict.go.jp/wn-ja/ (Japanese) Synset 05980875-n; queried 2014/01/15

These definitions might seem obvious. Yet, scientific usage of “goals” sometimes does go *against* this definition, and it is important to see how. Also, the first and third point leave a lot of space for interpretation. We discuss examples for both cases in this section. Of course, dictionary definitions are not a hard constraint for scientific usage, and vice versa. What we discuss here is a matter of definition, not definite right or wrong. Therefore we aim to work out these differences first, and develop a terminology that is suited to discuss and understand these differences within the interdisciplinary community.

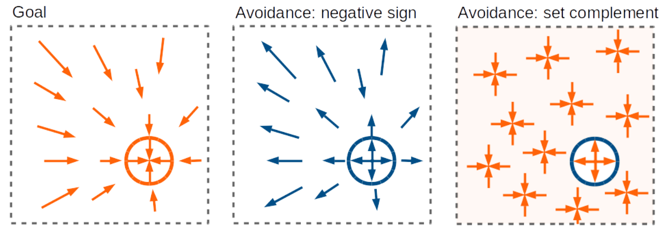

Fig. 1. (Left): Goals describe states towards which behavior is directed. (Middle): Just reverting behavior to describe avoidance leads to divergence. (Right): Avoidance can be described by all other states as target.

A. *Goals in BDI Architectures*

Goals have for a long time played an important role in artificial intelligence. For instance in *belief-desire-intention* (BDI) architectures [17], [18] goals are seen as instances of particular “desires”, while “intentions” describe the active, momentary pursuit of them. Out of this a large variety of particular meanings and implementations has evolved [8]. For instance, achieve goals describe the common sense of directing action towards some end-state. Some particular architectures (see [8]) have brought up maintain goals, which do not actually describe a goal in a different sense, but give them an extended life-cycle (e.g. staying active after the state has been initially reached). Similarly, test and query goals can be seen as desired states: they express the desire of the agent to know a certain fact [8]. Hence they refer to a desired internal state of the agent rather than an external one. These four types of goals are fully compatible with the above description.

Yet, there have been other kinds of goal implementations that do not fit the general definition. For example preserve goals are purely passive observation processes that monitor whether some condition is true. Braubach argued that these are “merely called a goal” [8], i.e. they are not really goals. In fact it seems plausible that over the history of BDI systems development various actions have been declared as “goals” within BDI only because there is no other way to introduce an action within these architectures. preserve goals describe an internal process that neither describes an actual desire, nor any directedness, but could not have been introduced into BDI if not making it one of the “goals”. This is very well visible for another kind of goal: perform. It describes a mere action, with no other end than the action being performed or not [8]. It is explicitly something *not* directed to any goal, but something that is just done. Hence, we suggest to keep these two cases of preserve and perform goals out of a general definition of goals. BDI architecture have been using another three kinds of goal implementations, which we will pick up in the remainder of this section.

B. *Avoidance Behavior*

A very different kind of behavioral aspect that has been considered as a goal already in the early 20th century by Lewin [19] is avoidance behavior [20]. Moskowitz described goals as world states that are either approached *or* avoided [21]. This notion can include the avoidance of simple obstacles or dangerous entities like predators or fire. In BDI architectures such behavior has been explicitly modeled by cease (as opposed to achieve) and avoid (as opposed to maintain) goals [8].

Obviously, this describes a relevant behavior. However, it is completely contradictory to the dictionary definition as well as common linguistic use. Pointing out that a child is trying to avoid a hotplate in the kitchen by saying “The child’s goal is the hotplate” does not work. It points out the exact opposite because “goal” has a naturally positive meaning of being directed *towards* something. Now, is avoidance a goal or not? Again, scientific usage of a word does not have to correspond to non-scientific usage.

The distinction of approach and avoidance could mean that simply the direction, or “sign” of behavior, is different. This would indicate the existence of two fundamentally different kinds of goals. When the sign is flipped in a goal-directed process, does it correspond to avoidance? We have depicted a very typical situation of approaching a goal (orange circle) in Fig. 1 (left). The agent would go towards the goal *from anywhere* else. Sometimes, but not necessarily, this is modeled by linear approach dynamics. If avoidance would indeed be a goal “in the opposite direction” (negative sign) the result would look like Fig. 1 (middle): behavior diverges from the circled area, and keeps diverging *everywhere*. We are not aware of a single case in which avoidance has actually been modeled like that. Rather, avoidance refers to a *local* behavior, in which some area is avoided, but which does not require further action if a sufficient distance is reached. Does that mean that avoidance behavior corresponds to a goal that is even more different from other goals than just by its sign? We suggest no. We suggest to think of avoidance, in fact, in terms of the same “approach” goals as previously considered – by considering *which* states are the desired ones. Fig. 1 (right) shows a set of desired states that comprises *all* world states *except* those in the avoided area (blue circle). This situation exactly describes avoidance behavior that is local and stops after a sufficient distance has been reached. This can be read as “the child’s goal is [not hotplate]”, instead of the hotplate itself being an anti- or negative goal. Hence, while the *behavior* of approaching or avoiding is directly opposed (relative to an object), we think that both can be understood in terms of the same positive goal definition. In this formulation avoidance does not refer to a totally different kind of goal, but to a particular way the set of (desired) world states looks like. In fact, already Lewin [19] was skeptical whether approach/avoid is a crucial distinction because already behavior resulting from different approach goals can take largely different forms.

C. Optimization and Reward

An interesting case of using the word goal is the context of reward optimization. Linguistic usage includes examples like "[The agent's] goal is to maximize future success or utility" [22] and variations of it [23]. Optimization also has an implementation as goal in some BDI architectures, where optimize refers to either the minimization or maximization of some variable [8]. Is this compatible with the desired end-state definition of goals? If not, what are the differences and should the preliminary definition be extended?

Obviously, reward is both about action and desire, which leaves the question how rewards relate to end-states. Acting towards improving one's reward is not *about* a *particular* state in the first place, but rather about finding *any* better one [16]. However, the ultimately accomplished optimization might be seen as to refer to the set of states that give a higher reward than any other state. In this line of argumentation, Montague noted that reward maximization might be seen as "highly abstracted definition of the goal" [24].

While goals are often considered as a very explicit set of world states (e.g. variables values or ranges, knowing whether something is true or not), we might also think of more implicit definitions. For instance, puzzles and riddles define criteria for a solution, without giving the solution right away. In the same way $n^2 = 4, n \in \mathbb{N}$ is an implicit formulation of $n = 2$. Hence, we might call reward optimization a goal referring to the set of states that give maximum reward.

There is, however, one difference between sets of maximum rewards and other set-of-states definitions such as in riddles and frequently employed meanings of goals. Many research papers have pointed out the distinct role of goals (in humans and/or AI) to determine whether a course of action was successful or not, and if not to be able to compare actual and desired state [25], [26], [2], [8], [27], [28]. This test of achievement is *not generally* possible for reward optimization. Determining the global optimum of a reward function is not generally possible for any kind of function – not even how much better one could get. Even for rather simple special classes (e.g. quadratic optimization) the problem is NP-hard [29]. Thus, in the general case an agent can by no means know whether he has actually reached the desired set of states.

As an example of this difference we can consider high jump athletes. We could say: "The goal of the athlete is to jump as high as possible." Yet, no one can for sure indicate what height this refers to. Is the current world record of $2.45m$ the maximum? We do not know whether or even by which margin it is not. Before the invention of the Fosbury Flop athletes optimized different techniques to jump higher and higher. Yet, they had to discover a completely new technique to leave the old local optimum and find a new, better local optimum. Still, the global optimum might not have been discovered.

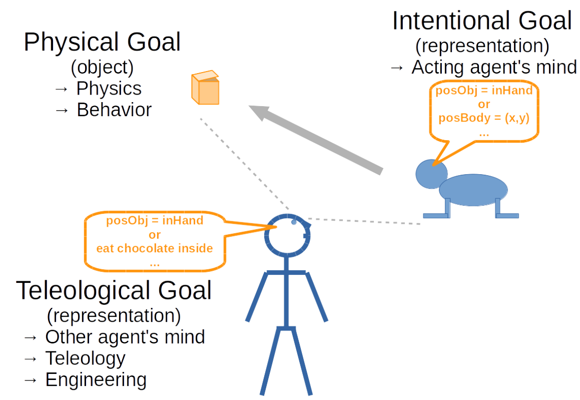

Fig. 2. Physical goals are actual objects towards which behavior is directed. Intentional goals and teleological goals are representations of such end-states in the mind of an acting and observing agent respectively.

Many authors have not only emphasized the need for explicit goals [25], [26], [2], [8], but also stated that goals and rewards are simply different: goals require knowledge and representation [30], [31], [1]. We will pick up the aspect of representation in the next section. For the purpose of this paper's discussion we agree with the latter authors to keep goals and rewards as separate concepts on a general level. Goals and rewards are, first of all, not the exact same thing. In the next section we will argue that, nevertheless, reinforcement learning agents can be immediately *understood in terms of goals*. We will then further concretize this relation in Sec. IV and V and discuss various ways to connect them.

III. 3 Perspectives on Goals

In the previous section we have argued that a vast majority of goal concepts can indeed be considered desired end states of action. In this section, we discuss different perspectives on exactly this matter that result in very different implications.

A. Physical Goals

The first very distinct perspective is that of *physical* goals. For instance, in common language use, physical "goals" occur in sports. In soccer and other sports the physical pair of posts with a cross bar is called "goal" in English language – so is the event of scoring. In running events the finish line is called "Ziel" ("goal") in German language.

This pattern can also be found in scientific literature. For instance, in animal behavioral studies "goal-tracking" can refer to the behavioral tracking of a physical goal [32] – an *object*. In particular, objects that are attributed with an inherent value "such as foods and fluids" [33] are frequently called goal. Some authors explicitly noted that "goal-direct behavior" in their view exclusively concerns behaviors that are oriented to external, physical objects [24]. It is very important to see that this perspective is very distinct from others: it does *not necessarily* involve an actual wish, desire or intention of an agent. It only concerns the physical world of objects and movement or behavior which is possibly directed towards that object. In the same way, the movement of a stone rolling down a slope into a valley might be described by means of the valley as a "goal" [34]. This is certainly a philosophical extreme position, but it correctly points out the pure physical character of this goal notion in the possible absence of any internal process corresponding to the physical events.

The same distinction between physical objects and events in the world versus internal processes within an agent has led to the technically correct distinction between "rewards" (objects) and "reward signals" (e.g. dopamine signals) in neuroscience. We suggest the same distinction here and call the objects from a physical world perspective "physical" goals as opposed to the other two perspectives we discuss in this section.

B. Intentional Goals

Another perspective – on possibly the exact same thing – is the internal operating of an acting agent. In this sense numerous authors from different fields have pointed out that goals in their sense refer to explicit, internal *representations* of what the agent is going for ([34], [30], [31], [1], [24], [18], [8], [21]). This has also been argued to be the key difference between agents with goals and merely reinforcement learning agents ([30], [31], [1]) that do not have explicit, "declarative" [25] knowledge about the ends of their action. "Goals reside in memory as mental representations" [21]. Such representations should in particular reflect what aspects of a situation are relevant in relation to action and the underlying desire [1] and thus abstract from irrelevant aspects ([26], [21]).

In the philosophy of mind this aspect is immediately linked to the concept of *intentionality* [35], [34]. Here intentionality (not be confused with common sense "intention") refers to "the power of minds to be about, to represent, or to stand for, things, properties and states of affairs" [7]. This notion of goals is also used in neuroscience within the concept of agents' "*cognitive control*" [24]. With smaller cognitive or philosophical meaning such explicit internal goal representations are also used in standard control systems (e.g. a robot's arm control) [36], [37] or standard planning systems [27]. In such systems goals describe which variables or properties of the environment or agent should have which value.

This kind of goal notion refers to the same kind of thing, i.e. end states, just from a different perspective. Yet, this difference in perspective causes largely different implications. Also, an intentional goal might differ from a physical goal in various ways when considering particular instances of behavior. Physical goals typically refer to a particular thing such as an apple, while the intentional goal could be food in general. In other cases, the agent might not be able to reach its intentional goal because of a lack of procedural skill. In a behavioral experiment, an object called "goal" by the experimenter might have entirely nothing to do with an agent's actual intentional goal. Other intentional goals might have no external physical correspondence, such as "query" goals in BDI [8] which refer to an internal state of knowledge.

C. Teleological Goals

We have discussed the physical view and the acting agent's internal view on goals. Another very common perspective is that of *other* agents *on* the acting agent. Imitation [38], [39] heavily involves this perspective as an agent tries to understand another one's action and utilize it for himself. Thereby the "interpreted" goal is not necessarily the same as the acting agent's intentional goal, and the observing agent itself might have a different intentional goal. In fact, research on imitation also involves an additional strong physical perspective when imitation is said to be "directed" to a physical goal [38], [39].

We use the word "teleological" goal in this paper because of the outside perspective, trying to make sense of something else. In this way, already infants pursue a teleological understanding of their environment [40], [41]. McFarland noted that humans in general seem to have a "teleological imperative" [42], forcing them to see the environment and other agents by attributing goals. In this way even behavior or motion that is only "apparently purposeful" [34], such as a stone rolling down into a valley, can be easily misinterpreted in terms of an intentional goal. Also scientific behavioral studies are easily misinterpreted in this way.

Teleological goals can be different from physical and intentional ones when talking about particular instances of behavior, yet all refer to the ends of action. Interestingly, teleological goals seem (just like intentional ones) to imply a *representation* – just in a different agent's mind. Our mind makes sense of others' action by linking them to our knowledge and our representations of the world. In this way research on mirror neuron systems seems to indicate a brain area in which our intentional goal and teleological goals (interpreting others' behavior) are commonly processed [6], [43] – and the neurons that do the actual representation are called mirror neurons.

In this paper we refer to "teleological goals" as representations used to explain actions that have not been self-generated. This includes the aforementioned understanding of observed action. The same concept could also be used within a single mind, for instance when actions were taken without deliberation and then are postdictively rationalized [44]. Interestingly, this thought can be smoothly extended to representations of "hypothetical" actions, for instance those an engineer has in mind for a robot to perform.

D. Goal vs. Reward: Revisited

We think that exactly this kind of teleological goals can often be observed for reinforcement learning agents that attempt to maximize their reward. Even if the agent does not have an intentional goal representation as discussed above, the engineer who builds the system can have a very clear representation – a teleological goal – of the agent's behavior in mind. With the vocabulary introduced so far we can indeed talk about goals of a reinforcement learning agent from at least two perspectives (see Fig. 3).

Firstly, an engineer would try to develop, or "shape" [45] a reward function that corresponds to his teleological goal of the agent's future behavior. In some cases – though not generally

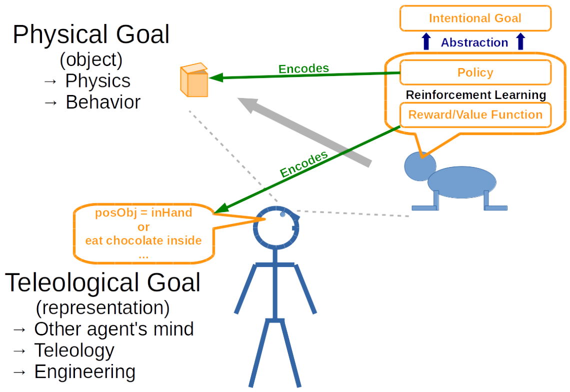

Fig. 3. In reinforcement learning, policies potentially encode physical goals among other things, while the reward/value function can contain a designer's teleological goals. Intentional goals might be learned as abstractions of those.

– this goal is coded explicitly into the reward function. For instance, the task to maneuver a robot arm through a via-point can be formulated by explicitly rewarding the distance to that point [46], or archery skills by rewarding the distance from a target [47]. Similarly, maze tasks typically come with explicitly encoded target states [23]. In this way the reward function encodes the designer's teleological goal. Then, *is* the reward function (or later "value" function) itself an intentional goal of the agent? We argue that it is not. Firstly because even if the teleological goal is explicitly contained in the reward function, it is typically not accessible to the agent but fully hidden inside the reward function. Secondly, and more importantly, the teleological goal might not be the only term in the reward function. Other factors such as energy or obstacle avoidance might play an equal role. Eventually, the agent's learning of a value function which predicts future reward removes any explicit appearance of the goal representation. Hence, both reward and value function *can contain* information about an explicit goal, but that is not the same as "being" a goal or directly representing it. After all, reinforcement learning agents can also have reward functions that do not contain an immediate teleological goal at all, but rather abstract criteria such as information gathering [48], [14], [16], [22]. In the same way, an agent's policy can correspond to physical goals by encoding the behavior directed towards it. This can also not be seen as a direct representation since the policy includes any other intermediate or transient part of behavior as well, and there is no explicit reference to the end-state.

This perspective on reinforcement learning also provides a description of how *inverse reinforcement learning* [49] is a teleological process. It starts with the physically observed behavior and estimates a reward and value function that can serve as a teleological explanation of that behavior. However, it does not provide an explicit teleological goal in the sense of a distinct representation of the actions' ends. We will argue in the next section how the learning of such explicit representations out of rewards could serve a possible explanation for the appearance of intentional goals.

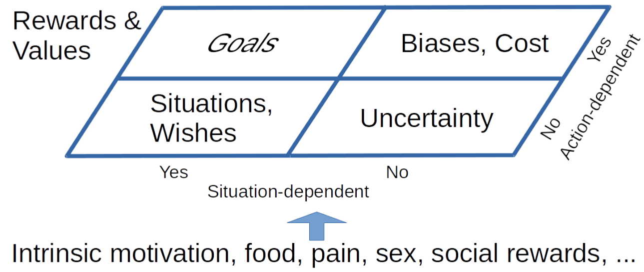

Fig. 4. Intentional goals might reflect rewards that are contingent with both the situation an agent is situated in as well as its actions [26].

IV. Do Rewards Originate Intentional Goals?

The common way in which nowadays AI systems have intentional goals is that an engineer immediately encodes them at the time of development. Humans, in contrast, are generally assumed *not* to come with pre-coded intentional goals right away – although they have them later in life as independent and individual agents. The process in between, that develops goals out of a state without them, is not generally clear. If the mind is believed to be essentially *deterministic*, there must be chains of events and mechanisms that form intentional goals. This does not have to be as "direct" as an engineer putting goals right into a system – a system that Dennett would call a "design stance" [35] in contrast to an "intentional stance". Yet, there must be some mechanisms. Such, possibly indirect or inexplicit, chains of mechanisms have sometimes been referred to as *origination* [50], [51], which is closely linked to debates about free will: "The will is the capacity of the self to originate goals which in turn determine the acts of the self" [51].

Thinking about the possible kind of such a mechanism vastly depends on how we define goals in the first place. We have initially pointed out that goals are an expression of desire. They have to reflect some value, which can be thought of in terms of internal reward signals. In fact, many authors have suggested an intimate function relation between goals and rewards [1], [30], [31], [14]. In some cases, this relation has been made explicit by stating that rewards could be the mechanism that originates intentional goals [24], [16]. In this section, we briefly review these proposals and discuss how this idea fits with the previously discussed relation of goals and rewards as well as with neuroscientific evidence.

A. Goals from Rewards

We had previously pointed out the idea to form intentional goals out of rewards [16], [26] based on an argument about possible information sources: if goals reflect a desire or value, then the information that originates goals should likewise reflect that value, at least in a more rudimentary manner. Starting from three possible learning signals in machine learning (supervised, unsupervised, reward), we argued that supervised learning involves an explicit ground truth signal that would already require the goal in advance. Unsupervised learning merely reflects signals statistics, which do not carry any inherent value. Reward signals, on the other hand, seem suited as a source of information as they come with simple value semantics pointing out how good or bad the current conduct is. From such reward signals an agent could learn representations of what concrete patterns within the potentially high-dimensional streams of sensory and motor information coincide with the reward. In this sense intentional goals could be seen as *abstractions* ([21], [26]) of action-related rewards describing what contingency of sensory context and action causes the reward (see Fig. 4). Also in hierarchical reinforcement learning, algorithms have been proposed that identify significant states towards which policies are directed (e.g. "bottleneck" states) in terms of goals ([52], [53], [54]).

Evidence and recent hypotheses in neuroscience point in the same direction. Brain areas that have long been associated with goals are the prefrontal cortex (in terms of decision making) [5] and the (pre-)motor cortex [55] (in terms of immediate action generation). Reward dependencies have been found in both parts of the cortex [56]. In the case of the prefrontal cortex, it has been explicitly hypothesized that reward-prediction errors, encoded by dopamine, trigger changes of goal-representation ("Dopamine gating hypothesis" [24]). This would indeed imply that at least prefrontal goal representations are formed out of a reward based learning signal as argued in [16], [26].

Such goal formation seems consistent with notions of goals or no-goals in conditioning scenarios [31], [30]. Instrumental conditioning is well formalized and understood in terms of reward-based reinforcement learning – yet not considered a highly cognitive or goal-driven behavior per se. Instrumental conditioning generates purely *procedural* knowledge – the "agent does not 'know' about the consequences of its behavior" [30]. Yet, Dickinson argued that such simple conditioning might turn into actual goal-directed behavior if (i) the goals are explicitly represented and (ii) correspond to the future expected reward and its contingency to the action [1]. While reward can have an immediate correspondence with teleological and physical goals, intentional goals might thus be the result of a learning process that abstracts reward relevant aspects into an explicit internal representation.

B. Goals not from Rewards

Intentional goals could be learned out of reward signals – supposing these are the basic manifestation of desire. While this seems in line with various conceptualizations of goals as well as neuroscientific evidence, we can not generally exclude the existence of other influences on human goal formation or the existence of goals being formed without any relation to rewards. In fact multiple authors have discussed a formation of goals out of mere sensorimotor contingencies without considering reward signals or intrinsic desires but in terms of anticipations described by ideomotor theory [57], [58]. Others have argued that goals are formed through imitation [4].

Goals that are not originated from rewards or values but from other, possibly supervised or unsupervised sources of information create two challenges: Firstly, it is not clear in what sense they could generally reflect a desire at all. Sensorimotor contingencies as well as observed action from others clearly point out *possibilities* ([59], [60]) of action. Yet, not all possibilities are generally desirable. Consequently, it is not clear *which* particular goal an agent should try to pursue in a given moment if that representation is not rooted in the agent's own desire. Secondly, they would describe a separate complete branch of action steering and decision making – one that by definition is distinct from reward-signal based decision making. If that would be the case we would need to find models and arguments for *selecting* one of both systems for a particular decision. Theories of cognition commonly assume that we *weigh* and compare our options (e.g. different instances of possible goals) in order to make a decision on what to do [17], [50], [51], [35], [34]. Such comparison could not be performed by means of reward (supposing internal reward signals also reflect a system's "value"), unless the goal without reward-origination is attributed post-hoc with a value through some loophole. If such post-hoc attribution does not take place, a general decision rule like "unsupervised goal always wins against reward-related goal" could take place. This would, however, describe situations in which we are not capable of weighing our options but we pursue a goal regardless of its value.

We are in no position to simply deny the existence of non-reward or value related goals. However, it is not clear how they could reflect desires or values which are commonly assumed to be the basis of our decision making [17], [50], [51], [35], [34]. Rather, they would seem to generate a situation of potentially contradicting, yet not comparable, decision systems – which is a quandary to the extent that it is unclear how a decision mechanism could look like.

V. Benefit of Intentional Goals over Rewards

So far we have discussed considering goals and rewards first of all as different things, but also that rewards can immediately reflect teleological goals and could be used to form intentional goals. In that case, intentional goals would in fact not point to anything else than what is expected to be rewarding anyway. Intentional goals and rewards would describe the exact same behavior. If that should be the case, then what could intentional goals be useful for? Considering them independent leaves the quandary of not knowing how to decide between them. Yet, could not an agent just stick to rewards without representing intentional goals? At first sight, reward could seem sufficient for a cognitive system. The interesting question that might tell us about the nature of goals and their role for cognition is: under which circumstances or assumptions could reward be enough; and how could intentional goals be useful if such conditions are not met.

We can consider a classical reinforcement learning system and – at this moment – assume that it has already learned a full model of its environment and reward. It could select an optimal action based on its value function, or exploit an already learned optimal policy. This would be fully sufficient – unless the agent's *computational resources* are too little to fully evaluate its model, e.g. in very high dimensions of continuous measures. In such computationally expensive domains the learning of representations such as dimensionality reduction or auto-encoders has long been considered. Goal representation expressing explicitly what is relevant for reward, compared to representations based on unsupervised signal statistics, could be very useful when facing such constraints. Early experimental evidence [26] supports that learned goals can save computational resources.

A. Quality as Source of Information

Besides facing such resource constraints, a system first faces the challenge to explore and learn. It is well known that goal-representations (*if* they already exist) can be useful for the acquisition of skills. This advantage is, for instance, exploited in goal babbling [15], [59]. In relation to reinforcement learning, it has been traditionally argued that while rewards might describe the more general setting, learning with goals and corresponding representations of action outcomes, as in motor control [37], could provide a more informative signal for learning and control [61]. Rewards only provide a scalar learning signal corresponding to "magnitude" [28] of how good/bad an action is. In contrast, explicit goal representations in relation to the outcome of an action provide "directional information" [28]: they also point out how far the goal has been achieved or not, and in which direction an improvement should be conducted [2]. An analogy of this is the relation between chemotaxis and visually approaching an object. Chemotaxis is delicate since it has to rely on a scalar measure of how close a physical goal could be. If the object is visually perceivable as goal, it is possible to translate this into an immediate direction in which an agent can move to get closer.

Not very much is known yet about situations in which such representations are *not* available right away. Research in hierarchical reinforcement learning suggests that the identification of "bottleneck" states as goals can indeed improve and accelerate ongoing learning [52], [53]. It seems plausible that also more general learning of goals as abstractions from rewards could guide and facilitate learning, but that still needs to be shown.

B. Modular Systems

A further possible advantage of intentional goal representations could be the integration of reward-based learning into larger cognitive architectures. Systems that are not monolithic reinforcement learners, but modular, have to exchange information between different modules in terms of some representation. For instance, authors in neuroscience have suggested analogies of parts of the brain and different learning mechanisms [62], [63]: unsupervised learning in the cortex, supervised learning in the cerebellum, and reinforcement learning in the midbrain. Reinforcement learning mechanisms could simply emit a copy of the expected reward signal, but this can not tell *what* is currently pursued. Value-functions or policies could tell, but it seems unlikely that a neural system could transmit them as entire functions. Intentional goal representations *of* the rewards, on the other hand, seem to be a suitable *interface* between reinforcement learning and (un-)supervised learning in a larger cognitive architecture. For instance, reinforcement learning parts could inform other structures *about* present desires and intentions, that could then communicate them by means of speech, or relate the own goals to observations from imitation learning.

VI. Conclusions

What are goals? And if so, how many? There is no inherently right answer to these question. Rather, we need to define such terms in a way that is useful for scientific communication. We have reviewed a wide variety of terminologies and argued that desired end-states of action are at the core of the vast majority. Still, there can be very different implementations, and in particular very different perspectives on goals. Within a larger interdisciplinary research context it is therefore crucial to understand these perspectives in order to understand other researchers' arguments and draw the right conclusions from experiments.

Notions of rewards and goals have often been used in the same context, yet sometimes synonymously and sometimes as distinct concepts. We have therefore proposed to look at goals from three different perspectives, each of which can be related to rewards. In particular we discussed previously proposed connections between rewards and intentional goal representations corresponding to such rewards. We have speculated about and discussed evidence in support of the utility of such goals in a learning system. Yet, the role of such goals in developing systems is far from being completely understood and we hope that the discussions in this paper can further facilitate research on it.

Acknowledgments

This study has been supported by the JSPS Grant-in-Aid for Specially promoted Research (No. 24000012). This research is related to the European Project CODEFROR (PIRSES-2013-612555).

References

[1] A. Dickinson and B. Balleine, "Motivational control of goal-directed action," *Animal Learning & Behavior*, vol. 22, no. 1, pp. 1–18, 1994.
[2] B. H. Ross, Ed., *The Psychology of Learning and Motivation: Advances in Research and Theory*. Academic Press, 2011, vol. 54.
[3] E. A. Locke and G. P. Latham, "Goal setting theory," *Motivation: Theory and research*, pp. 13–29, 1994.
[4] C. Heyes, "Causes and consequences of imitation," *Trends in cognitive sciences*, vol. 5, no. 6, pp. 253–261, 2001.
[5] S.-I. Kim, "Neuroscientific model of motivational process," *Frontiers in psychology*, vol. 4, 2013.
[6] G. Rizzolatti and L. Craighero, "The mirror-neuron system," *Annu. Rev. Neurosci*, vol. 27, pp. 169–192, 2004.
[7] P. Jacob, "Intentionality," in *The Stanford Encyclopedia of Philosophy (Winter 2014 Edition)*, E. N. Zalta, Ed., 2014. [Online]. Available: http://plato.stanford.edu/archives/win2014/entries/intentionality/
[8] L. Braubach, A. Pokahr, D. Moldt, and W. Lamersdorf, "Goal representation for BDI agent systems," in *Programming Multi-Agent Systems*, ser. Lecture Notes in Computer Science. Springer, 2005, vol. 3346, pp. 44–65.
[9] J. Weng, J. McClelland, A. Pentland, O. Sporns, I. Stockman, M. Sur, and E. Thelen, "Autonomous mental development by robots and animals," *Science*, vol. 291, pp. 599–600, 2001.
[10] J. Piaget, *The Origin of Intelligence in the Child*. Routledge and Kegan Paul, 1953.

[11] B. I. Bertenthal, "Origins and early development of perception, action, and representation," *Annual Reviews Psychology*, vol. 47, pp. 431–459, 1996.
[12] C. von Hofsten, "Eye-hand coordination in the newborn," *Developmental Psychology*, vol. 18, no. 3, pp. 450–461, 1982.
[13] A. van der Meer, "Keeping the arm in the limelight: Advanced visual control of arm movements in neonates," *European Journal of Paediatric Neurology*, vol. 1, no. 4, pp. 103–108, 1997.
[14] P.-Y. Oudeyer, F. Kaplan, and V. V. Hafner, "Intrinsic motivation systems for autonomous mental development," *IEEE Trans. Evolutionary Computation*, vol. 11, no. 2, pp. 265–286, 2007.
[15] M. Rolf, J. J. Steil, and M. Gienger, "Goal babbling permits direct learning of inverse kinematics," *IEEE Trans. Autonomous Mental Development*, vol. 2, no. 3, 2010.
[16] M. Rolf and M. Asada, "Autonomous development of goals: From generic rewards to goal and self detection," in *IEEE Int. Joint Conf. Development and Learning and Epigenetic Robotics (ICDL-EpiRob)*, 2014.
[17] M. Bratman, Ed., *Intention, Plans, and Practical Reason*. Harvard University Press, 1987.
[18] A. Haddadi and K. Sundermeyer, "Belief-desire-intention agent architectures," in *Foundations of distributed artificial intelligence*, 1996, pp. 169–185.
[19] K. Lewin, "Intention, will and need," in *Organization and pathology of thought: Selected sources*, D. Rapaport, Ed. Columbia University Press, 1951, pp. 95–153.
[20] K. Fujita and K. E. MacGregor, "Basic goal distinctions," in *Goal-Directed Behavior*, 2012, pp. 85–114.
[21] G. B. Moskowitz, "The representation and regulation of goals," in *Goal-Directed Behavior*, 2012, pp. 1–47.
[22] J. Schmidhuber, "Driven by compression progress: A simple principle explains essential aspects of subjective beauty, novelty, surprise, interestingness, attention, curiosity, creativity, art, science, music, jokes," in *Anticipatory Behavior in Adaptive Learning Systems*. Springer Berlin Heidelberg, 2009, pp. 48–76.
[23] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. MIT Press, 1998.
[24] P. R. Montague, S. E. Hyman, and J. D. Cohen, "Computational roles for dopamine in behavioural control," *Nature*, vol. 431, pp. 760–767, 2004.
[25] M. Winikoff, J. Harland, and L. Padgham, "Linking agent concepts and methodology with can," in *Int. Joint Conf. Autonomous Agents and Multi-Agent Systems (AAMAS)*, 2002.
[26] M. Rolf and M. Asada, "Where do goals come from? A generic approach to autonomous goal-system development," 2014, (submitted). [Online]. Available: http://arxiv.org/abs/1410.5557
[27] J. Bruce and M. Veloso, "Real-time randomized path planning for robot navigation," in *IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, vol. 3. IEEE, 2002, pp. 2383–2388.
[28] A. G. Barto, "Reinforcement learning in motor control," in *Handbook of Brain Theory and Neural Networks*. Cambridge: MIT Press, 1994, pp. 809–813.
[29] S. Vavasis, "Complexity issues in global optimization: A survey," in *Handbook of Global Optimization*, ser. Nonconvex Optimization and Its Applications, R. Horst and P. Pardalos, Eds. Springer US, 1995, vol. 2, pp. 27–41.
[30] B. W. Balleine and A. Dickinson, "Goal-directed instrumental action: contingency and incentive learning and their cortical substrates," *Neuropharmacology*, vol. 37, no. 4, pp. 407–419, 1998.
[31] W. Schultz, "Behavioral theories and the neurophysiology of reward," *Annu. Rev. Psychol.*, vol. 57, pp. 87–115, 2006.
[32] S. Kaveri and H. Nakahara, "Dual reward prediction components yield pavlovian sign- and goal-tracking," *PLOS one*, vol. 9, no. 10, 2014.
[33] B. W. Balleine, "Sensation, incentive learning, and the motivational control of goal-directed action," in *Neurobiology of Sensation and Reward*, G. JA, Ed. CRC Press, 2011, ch. 13.
[34] D. McFarland, "Opportunity versus goals in robots, animals and people," in *Comparative approaches to cognitive science*. MIT Press Cambridge, MA, 1995, pp. 415–433.
[35] D. C. Dennett, Ed., *The intentional stance*. MIT Press, 1989.
[36] W. Chung, L.-C. Fu, and S.-H. Hsu, "Chapter 6: Motion control," in *Handbook of Robotics*, B. Siciliano and O. Khatib, Eds. Springer New York, 2007, pp. 133–160.
[37] D. Nguyen-Tuong and J. Peters, "Model learning for robot control: a survey," *Cognitive Processing*, vol. 12, no. 4, 2011.
[38] H. Bekkering, A. Wohlschlager, and M. Gattis, "Imitation of gestures in children is goal-directed," *The Quarterly Journal of Experimental Psychology: Section A*, vol. 53, no. 1, pp. 153–164, 2000.
[39] A.-L. Vollmer, M. Mühlig, J. J. Steil, K. Pitsch, J. Fritsch, K. J. Rohlfing, and B. Wrede, "Robots show us how to teach them: Feedback from robots shapes tutoring behavior during action learning," *PloS one*, vol. 9, no. 3, 2014.
[40] G. Csibra, S. Biro, O. Koos, and G. Gergely, "One-year-old infants use teleological representations of actions productively," *Cognitive Science*, vol. 27, pp. 111–133, 2003.
[41] G. Gergely and G. Csibra, "Teleological reasoning in infancy: the naive theory of rational action," *Trends Cog. Science*, vol. 7, no. 7, pp. 287–292, 2003.
[42] D. McFarland, "The teleological imperative," in *Goals, No Goals and Own Goals*, Montefiore and Noble, Eds., 1989.
[43] M. Iacoboni, I. Molnar-Szakacs, V. Gallese, G. Buccino, J. C. Mazziotta, and G. Rizzolatti, "Grasping the intentions of others with one's own mirror neuron system," *PLoS biology*, vol. 3, no. 3, 2005.
[44] S. Shimojo, "Postdiction: its implications on visual awareness, hindsight, and sense of agency," *Frontiers Psychology*, vol. 5, no. 196, 2014.
[45] A. Y. Ng, D. Harada, and S. Russell, "Policy invariance under reward transformations: Theory and application to reward shaping," in *Int. Conf. Machine Learning (ICML)*, 1999.
[46] E. Theodorou, J. Buchli, and S. Schaal, "Reinforcement learning of motor skills in high dimensions: A path integral approach," in *IEEE Int. Conf. Robotics and Automation*, 2010.
[47] P. Kormushev, S. Calinon, R. Saegusa, and G. Metta, "Learning the skill of archery by a humanoid robot icub," in *IEEE-RAS Int. Conf. Humanoid Robots*, 2010.
[48] J. Gottlieb, "Attention, learning, and the value of information," *Neuron*, vol. 76, no. 2, pp. 281–295, 2012.
[49] A. Y. Ng, S. J. Russell *et al.*, "Algorithms for inverse reinforcement learning," in *ICML*, 2000, pp. 663–670.
[50] T. Honderich, Ed., *A theory of determinism: The mind, neuroscience, and life-hopes*. Oxford University Press, 1988.
[51] M. D. Robertson, "Dualism vs. materialism: A response to paul churchland," 1995.
[52] I. Menache, S. Mannor, and N. Shimkin, "Q-cut—dynamic discovery of sub-goals in reinforcement learning," in *European Conf. Machine Learning (ECML)*, 2002.
[53] O. Simsek, A. P. Wolfe, and A. G. Barto, "Identifying useful subgoals in reinforcement learning by local graph partitioning," in *Int. Conf. Machine Learning (ICML)*, 2005.
[54] M. van Otterlo, Ed., *The logic of adaptive behavior: knowledge representation and algorithms for adaptive sequential decision making under uncertainty in first-order and relational domains*. Ios Press, 2009.
[55] B. Pesaran, M. J. Nelson, and R. A. Andersen, "Dorsal premotor neurons encode the relative position of the hand, eye, and goal during reach planning," *Neuron*, vol. 51, no. 1, pp. 125–134, 2006.
[56] M. R. Roesch and C. R. Olson, "Impact of expected reward on neuronal activity in prefrontal cortex, frontal and supplementary eye fields and premotor cortex," *Journal of Neurophysiology*, vol. 90, no. 3, pp. 1766–1789, 2003.
[57] S. Verschoor, M. Weidema, S. Biro, and B. Hommel, "Where do action goals come from? evidence for spontaneous action-effect binding in infants," *Frontiers in Psychology*, vol. 1, no. 201, pp. 1–6, 2009.
[58] A. Greenwald, "Sensory feedback mechanisms in performance control: With special reference to the ideo-motor mechanism," *Psychological Review*, vol. 77, pp. 73–99, 1970.
[59] A. Baranes and P.-Y. Oudeyer, "Active learning of inverse models with intrinsically motivated goal exploration in robots," *Robotics and Autonomous Systems*, vol. 61, no. 1, pp. 49–73, 2013.
[60] M. Rolf, "Goal babbling with unknown ranges: A direction-sampling approach," in *IEEE Int. Joint Conf. Development and Learning and Epigenetic Robotics (ICDL-EpiRob)*, 2013.
[61] L. P. Kaelbling, M. L. Littman, and A. W. Moore, "Reinforcement learning: A survey," *arXiv preprint cs/9605103*, 1996.
[62] N. D. Daw and K. Doya, "The computational neurobiology of learning and reward," *Current Opinion in Neurobiology*, vol. 16, no. 2, 2006.
[63] W. Schultz, P. Dayan, and P. R. Montague, "A neural substrate of prediction and reward," *Science*, vol. 275, no. 5306, 1997.