Scaletta Power point

The presentation goal is to demonstrate the following abilities:

a. Communicate the key idea of the article

We need cooperative agents who really understand people.

Discorso sulla difficoltà a capire e sulla insufficienza per capire una persona di solo quello che è logico/il suo utile.

Why question

What if

* Cooperative agents
  + Why
  + What if

b. Illustrate the novel technical aspects, or the proposed methodology, if any

c. Analyze critically the proposed ideas: is the line of argumentation based on solid assumptions? Is there a hidden agenda? Can we agree with the article assumptions?

d. Analyze critically the methodology: is an evaluation of the method presented? Is it valid? (i.e., does it evaluate what needs to be evaluated?)

e. Analyze critically the communication: is the language clear? Is the organization adequate? Are examples provided?

Linguaggio poco chiaro anzi quasi per nulla. Lessico complesso e pesante.

Gli esempi sono presenti, non sufficiente però per capire molte cose e soprattutto alcuni esempi sono scritti coi piedi.

Organizzazione ok, perlomeno si procede a step.

f. Summary of the strengths and weaknesses of the paper

Strengths: Interessante e stimolante per la discussione fra di noi

Weakness:Pesantissimo da comprendere, inglese non fluido

g. What have you learnt from this article.

Nemmeno le persone reali capiscono le altre figuriamoci sta roba. Potrebbero capire e prevedere un comportamento ma non sempre e soprattutto troppe variabili da tenere conto.

Neanche noi ci conosciamo

Ma poi una macchina che generalizza come può adeguarsi alla variabilità degli esseri umani? Ci sono assassini e persone tranquillissime, hanno optmal policies diversissime.

Si parla di un box in più occasioni e poi lo si mette 4 pagine dopo.

Inglese pessimo: If a cooperative AI is to explain human behavior, then it must do so in terms of a model of these components of the problem and a derived optimal policy.

Come si matematizza il pensiero di un umano? Va stimato anche il suo non essere razionale.

Come si fa a capire in base a cosa sono misinterpretati? Magari ha ragione Kahneman, manco lo spiega perchè: In cognitive science,it has been shown that explanations for some puzzling aspects of human behavior have some times been misattributed amongs these three components of the bounded optimization problem. Kahneman for example, attributes a number of human behaviors to distorted utility functions when these should, instead, be attributed to the structure of the task environment.

Punti

Computationally Rational model: explains human behavior by establishing a causal connection between subjective utility, capacities, and experience on the one-hand and behavior on the other. In other words, cooperative AI needs a computational model of the human that is defined in terms of an objective function representing human subjective utility, a model of human capacities (memory, perception, reasoning, etc.), and a model of how humans experience the task environment over a lifetime. In addition, to be utilized in cooperative AI, the general human model needs to be “fit” to each individual collaborator. People vary widely along all three of the key dimensions. For example, in subjective utility, while most people are risk-averse, there is considerable variation (Mata et al. 2018), with some individuals even being risk seeking in the domain of gains. Un modello deve sapere ste preferenze, sta per avere a che fare con un sottone o un giga-chad?

While admitting to the possibility of variation in subjective weightings, computational rationality commits to a normative model of subjective utility, which is a requirement of the rationality assumption. Rather than appealing to “biases” in choice, a computationally rational account explains behavior as an emergent consequence of resource limits rather than as a consequence of “irrational” policies

In many, though not all, computationally rational models, the optimal policy is generated by defining the bounded optimality problem as a Markovian decision problem (MDP) and then solving it using classical planning algorithms. All of these problems share the fact that they can be used to precisely capture a sequential bounded optimalisation problem in which goals are expressed as cumulative reward maximization

Standard use di Markovian: agente interagisce con un ambiente esterno.

Computally rational model alternativi: agente interagisce con un ambiente interno con osservazioni e azioni con uno yoked(aggiogato) external environment via stimuli e responses.

In addition, people can make strategic choices about which task to prioritize and whether to pri oritize speed or accuracy. The fact that task performanceis influenced bybothcapacity limits and strategic concerns makes it difficult for an observer to tease out the contribution of each

La gente crede che gli altri siano razionali. Considerazione poco sensata.

Nella cooperative AI bisogna massimizzare una value function V che riassume i vantaggi di una interazione, al tempo stesso il partner umano deve massimizzare la propria utility Function U. Ergo gli obiettivi possono non essere gli stessi.

A task history ℎ∈H is generated from the model 𝑀, following a policy 𝜋 with parameters 𝜃 of the mechanisms of human cognition, in a task environment 𝜙. A computationally rational agent deploys a bounded optimal policy 𝜋∗ that maximizes the utility of the task: 𝜋∗=argmax 𝜋 V(ℎ), con h=M(pi,theta,task\_env)

Why?questions:Parameter inference determines the most plausible set of parameters𝜃∗ by maximizing the likelihood of observed human data 𝑦: 𝜃∗=argmax 𝜃 𝑝(𝑦∣𝑀). The likelihood is computed based on the model 𝑀 of the user interacting with the task environment 𝜙.

Whatif?questions:The agent attempts to optimize a set of interventions 𝑖∗⊂D(often called “designs”), that out of the space of all possible designs maximize the expected value function V,given the predicted history of behavioral sequences that are adaptations to these interventions by the user.

Recently, applications have emerged in two areas: (1) computational design and (2) adaptive user interfaces (UIs). In computational design, computationally rational models have been used to determine design objectives.

IRL per rispondere al WHY: the reward function of an observed agent is inferred from observations of its behavior under the assumption that the behavior was generated with an optimal policy (Equation 2). A typical IRL algorithm takes as input a set of behavioral trajectories and outputs a reward function.

The computational costs of solving the problems tend to grow rapidly with the size of the problem. More importantly, from the perspective of the current paper, IRL is not developed to take into account latent factors other than the reward function; factors that are critical to modeling humans including cognitive limitations and experienced ecology, which are known causal determinants of behavior, that vary across people. Moreover it assumes that the human policy is optimal with respect to the task at hand, which precludes the possibility that the human is optimal with respect to say, teaching the task; in other words, the human is performing the task in a way to make it easier for the IRL agent to learn. The human might position a hand in such a way so as to make an action visible, rather than so as to minimize movement time

Alternativa all’IRL per respondere al why è la likelihood free inference(FLI). In LFI, observations are compared to sampled predic tions made using the simulator. This permits approximat ing the plausibility of parameter values without an explicit likelihood function. Although the likelihood-function is not known, it does exist, and drawing a large number of samples from the model can give sufficient informa tion about the plausibility of different parameter values.Però con LFI troppa high-dimensionality ergo si reduce la dimensionalità creando summary statistic, la cui scelta e design è fondamentale per una inferenza di successo. Inoltre serve domain-specific knowledge, soprattutto se riduciamo la dimensionalità. By running the simulator multiple times and computing the difference between the summary statistics from the observed data and generated data, one can arrive at an approximation of the likelihood of different parameter values.

Because answers to “why?” questions should be expressed in terms of how likely different causes of observed behavior are, parameter inference techniques that estimate a posterior are preferable to those that estimate a fixed point.

LFI was recently used to understand task interleaving (Gebhardt, Oulasvirta, and Hilliges 2020). It explains why some people are more proficient at making efficient task switching decisions by modeling task interleaving as a hierarchical RL problem,where each available task is associated with its own cost and reward estimates.

Fitted to almost 200 individual participant data, the analysis suggested that just two parameters— the discount factor of the supervisory controller and a coefficient that controlled a trade-off between rewards and task switching costs— could explain much of individual differences observed; thereby answering a “why?” question. Users with high discount factor persisted longer in a task before switching to another task

Combining computationally rational models with LFI significantly enhances the power of both for explaining human behavior.

Humans also vary in capabilities and experience of the task environment.

Simon pointed out that it is impossible to understand human behavior without understanding that it is adapted to both the structure of the environment and cognitive limitations.

It’s impossible to understand human behavior without taking utility, capabilities, and experienced environment into account simultaneously.

The model is fit to individuals via parameter inference, which is carried out in Equation (2) and results in a set of parameters 𝜃∗ that govern the functioning of the model.

Computationally rational models achieve this by finding a policy that maximizes utility under bounds.

What computational rationality adds is that predictions are made by calculating the bounded optimal policy in a given environment.

Computationally rational models can explain why behavior differs among individuals and conditions, as opposed to just describing them.

Others researchers have gone as far as arguing that progress in AI will curtail without better use of models

Arguably, the latent states and processes that are characteristic of human cog nition are not captured easily by model-free methods that learn from patterns in observed behavior.

One core insight has been that the adaptive problems faced by people can be defined as reinforcement learning (RL) problems and solved accordingly.

The assumption that human behavior is boundedly optimal is controversial in cognitive science. However, there is evidence that people appear optimal when the bounds imposed by computation or environment are taken into account

For the purposes of cooperative AI, the question is not whether people are able to optimize behavior,but whether behavior can be predicted by an optimization algorithm.

Computationally, rational models view cognition as a bounded optimization problem

Bounded optimality is distinguished from bounded rationality, where, in the latter, the role of optimization in understanding behavior is rejected.

An agent is bounded optimal if its program is a solution to the optimization problem presented by its utility function (objective), capacity limits (architecture), and the task environment.

When a model of a human and the environment defines a bounded optimality problem, the solution is a program 𝑃∗ , a policy 𝜋∗.The program simulates the user’s strategy, which determines the choice of actions. 𝑃∗ is adapted through optimization to the bounds imposed by the tuple, possibly using machine learning. This assumption is highlighted by stating that the optimal policy 𝜋∗ is a function of the agent’s subjective utility, capacity limits, and experience of the environment (its history).

Computational rationality is a departure from the earlier idea that human behavior can be predicted by considering only the environment, and, “not the assumed structure of the mind”. It is also a departure from the idea that behavior is shaped to external rewards. Internal bounds matter.

Creare il model

1. Specify the external environment of the user.

2. Specify the internal environment as an MDP (Markovian decision process). It describes the user’s cognitive processes and the reward function.

3. Estimate cognitive parameters from human data

As the working example, we use a driver model created for intelligent lane assistance in semi-autonomous vehi cles. In intelligent lane assistance, the ideal timing and degree of assistance depend both on the traffic situation and the driver’s state.

* The first step of the workflow is to formalize the external environment. This is a radical departure from data-driven 10 AI MAGAZINE approaches, which usually begin with acquiring training data. The definition of the external environment involves modeling the states and transition dynamics of the exter nal environment, and how they are affected by the actions of both the human user and the computer agent. The transition function can be modeled as hand-coded state transitions or via a physics simulator. In the driving model, the external environment (𝜙) would contain those aspects of the cockpit and traffic environment that are important for lane keeping, such as the road, signs, relative positions and speeds of vehicles, and the dynamics of the traffic flow. Livello di dettaglio del modello dipende dalla difficolta dell’obiettivo di modelling.
* The second step is to specify the internal environment, which describes the user’s cognitive processes and states. The internal environment is specified as an MDP. This step involves operationalizing psychological hypotheses about how the human mind processes task-relevant information. This requires answering questions such as how human senses provide information, how this information is integrated as goal-directed mental representations, and how those representations are stored, processed, and retrieved when making decisions. In the driver model, the internal environment simulates how a human driver perceives the external environment, and uses “percepts” to form an internal representation of the state. The representation integrates the perceptual information with prior representations of the task using an internal model of the task environment. To account for more complex situations in driving, we may also need to model beliefs about the goals and behaviors of other drivers and a model of naive physics, or how drivers understand physics related to driving. Furthermore, beliefs such as how the driver represents the internal workings of the driving automation, are encoded here. Parameters 𝜃 are used to encode individual differences related to the internal environment; they can, for example, capture how novice versus expert drivers mentally represent and simulate the environment. As part of this step, one needs to specify the reward function. While humans have complex, multifaceted preferences and life goals, the modeler should consider mainly those goals that are relevant to the task at hand. In the driver model, the reward function quantifies the driver’s aim to reach the intended destination safely and efficiently. In addition, a driver may have preferences about how they can best enjoy the trip, including music or the choice of alternative routes. These preferences will often be in conflict; for instance, the driver may need to consult an in-car interface for navigation or selection of the song, but this results in glances out of road and reduces safety. The resulting reward function should express how a driver might trade-off such objectives. Such preferences are encoded as weight parameters included in 𝜃. Given the full model, the agent is trained using RL. It is trained through interaction with the model of the external environment. In the computational rationality workflow, unlike in earlier cognitive architecture models, the modeler does not need to predefine the policy. Instead, it is assumed that the agent’s policy is optimal (𝜋∗) within its bounds. The parameters 𝜃 are initially provided with a prior based on the modeler’s assumptions about the values that researchers may have. At this stage in the workflow, the model is complete and initial steps can be taken to check that the policy converges to reasonable behaviors for sampled values of the parameters
* The third step in the workflow is parameter estima tion. For instance, there may be a need to fit the model parameters 𝜃 to each individual human driver, or to a particular class of drivers. Both the internal and external environments have parameters that govern variation in the task and user behavior. Usually, the parameters of the external environment are well-known, but those of the internal environment are not. For example, the goals of the driver, how they process information and what knowledge they have, are unknown and must be inferred from their individual performance data. Per facilitare la stima dei parametri, the modeler can use a theoretical or empirical prior about plausible parameter values in the user population. However, as more observational data about a particular user are col lected, cooperative AI must perform parameter inference in order to establish the most likely parameter estimates for that user. Si va dal generic al particolare, di quella persona. Lei devo aiutare.

How could a model built with this workflow be tested and deployed? The first part of the answer to this ques tion, is very carefully and responsibly, taking full account of the potential dangers, and ensuring oversight at every step. Technically, the goal of a cooperative AI system is to design an optimal set of interventions 𝑖∗ ⊂  to maximize the . Practically, this happens by simulating the parameterized model and assessing what sorts of interactions maximize this value.For example it could boost the lane assist to help the driver to keep the lane, or provide the driver with information about the best lane, given where the driver is going.

Computational rationality is a model-based, first principles approach to understanding humans. We believe that it has the potential, in the long-term, to go beyond data-driven methods—thanks to combining machine learning with psychological assumptions about the data-generating process, the human. Critical to computational rationality is the assumption of rationality. Predictions about behavior are derived from a policy that is optimally adapted to utility and the bounds—possible causes of behavior.

The causal link, made available by the optimization assumption, makes it possible to use the model to answer the “what if?” and “why?” questions that are essential to the evaluation of interventions as well as to updating the model when new evidence is observed.

Many significant challenges remain if this vision is to be fully realized, however. One of the most significant concerns the nature of human motivation: What makes a person pursue some activities and not others? What, for example, makes one person give up in the face of adversity, while others persist? One part of the answer is, as we have argued, that explanations of behavior must be a joint function of reward function, resource limits, and environment. But another answer rests on a deeper understanding of the human reward function itself.

A rational agent should be motivated to explore stimuli that maximally increase the usefulness of its knowledge. Curiosity is studied as a mechanism by which humans approximate this rational behavior.

Another challenge concerns the human ability to gen eralize knowledge and skills- Without understanding this, AI will grossly underestimate human abilities in novel encounters, such as when facing an intervention the AI has taken. So far, most models of computational rationality use model-free RL, which does not perform well compared to model-based alternatives when the envi ronment changes and transfer of knowledge and skills is required.

Computational rationality may also appear to be at odds with human emotions. However, this perceived conflict arises only if one associates emotions with maladaptive and irrational behavior, and contrasts them with rational goal-directed adaptive behavior.

Finally, we believe that computational rationality is a candidate technology on the path to human–AI align ment. IRL and cooperative IRL will prove more effective if complemented with viable models not only of human utilities but also of human bounds. Rather than assuming that humans can be modeled as rational agents with varying reward functions, as in IRL, computational rationality also embraces variation in internal bounds, including variation in observation functions, internal state transitions, and environments of adaptation. The resulting optimization problems are more challenging but must be faced if IRL is to produce useful solutions to alignment problems.

We believe that model-free learning is insufficient for cooperative AI. Computational rationality offers a principled, model-based basis for algorithms to drive both inference and planning in cooperative agents. It puts psychological constructs, similar to those that underlie human cooperative abilities, at the center of such algorithms. This enables the theory to disentangle causal contributions that a person’s goals, capabilities, and environment make to their behavior. It offers an exciting avenue for research on cooperative agents that better understand humans, plan interventions, and make available explanations that are human-understandable.