Reviewer 1

This paper proposes a new method to solve deterministic qualitative Dec-POMDPs (QDec-POMDPs). The approach breaks the problem into a sequence of single agent problems and solves them, backtracking as necessary. The details of the method are given and experiments are provided, comparing to Dec-POMDP methods.

Dec-POMDPs and QDec-POMDPs are general models for decentralized multi-agent decision-making, but as the authors point out, scalability of current methods can still be an issue. Developing methods that are scalable, while also providing high-quality solutions is therefore an important challenge.

The general idea (transform the multi-agent QDec-POMDP problem into a series of single-agent problems, solving each one of them while making sure the solutions are consistent) is very reasonable. As pointed out by the authors, similar ideas have been tried in the case of Dec-POMDPs. In fact, there are are Dec-POMDP methods that are even closer in spirit to the proposed method (e.g., the one below)

Varakantham, Pradeep, et al. "Exploiting Coordination Locales in Distributed POMDPs via Social Model Shaping." ICAPS. 2009.

A: we will cite

But of course, the QDec-POMDP case is different and developing techniques in this context is novel.

The algorithm itself is somewhat straightforward (making optimistic assumptions about other agents and backtracking until solutions are consistent), but, again, novel in this framework. Getting this approach to work in QDec-POMDPs requires extensions such as determining and propagating the collaborative constraints. The authors also make the results more efficient after solution with their sub-goal assignment method. It is also beneficial that the approach allows use of efficient, off-the-shelf contingent planners.

A: positive point

The experiments show the method can solve large problems. The approach is more scalable than the Dec-POMDP methods on the box-pushing problem and also solves larger rover domains. As the authors point out, the comparisons are a bit unfair since the Dec-POMDP methods have a different objective (optimizing rather than achieving a goal). Also, why comparison with these Dec-POMDP solvers? There are many Dec-POMDP solvers. The most scalable optimal Dec-POMDP solver is probably the one below.

A: we consulted with leading Dec-POMDP experts (Amato, Spaan, Oliehoek) and they suggested these algorithms.

Dibangoye, Jilles Steeve, et al. "Optimally solving Dec-POMDPs as continuous-state MDPs." Journal of Artificial Intelligence Research 55 (2016): 443-497.

Similarly, DICEPS is probably one of the more scalable "approximate" Dec-POMDP methods, but many others exist. The authors should motive why the selected methods were used and comparisons with other algorithms are not included.

A: we will be happy to compare to more methods. As long as they support Cassandra (Dec) format this is very simple to do.

Also, why are comparisons for Dec-POMDP algorithms not included in Table 3 and 4? While the domains have large state spaces, the other problem variables are small (maxing out at 5 agents, 8 actions, 3 observations and 3 agents, 15 actions and 5 observations). Since the action and observation space are small, methods such as DICEPS should be applicable since they search directly in policy space (and policies shouldn't be too large). Infinite-horizon Dec-POMDP methods (which often represent the solution as a set of finite-state controllers rather than trees) may be more scalable as the problem horizon grows. The authors should discuss the domains and comparisons in more detail.

A: we tried running the algorithms on the problems from Table 3 (Table 4 is just an example of execution for IMAP) and they all failed (crashed).

The quality of the results is also not discussed. It is hard to know what a "good" solution is here. The goal is reached, but the expected cost is much higher for the proposed method in Table 2 and beyond the small example at the end of 4, what the solutions are like is unclear. An expanded discussion would be helpful. More details would be also be beneficial on which classes of problems the authors expect sound solutions could be generated with their method.

A: we will try to squeeze some discussion in concerning the quality of solution. A major source for non-optimality in IMAP is the underlying classical planner (FF), which seems to be making sub-optimal choices. We are investigating moving to an optimal classical planner.

The writing is generally clear. It is beneficial that they use the box-pushing example, but it would be better to include a figure and additional details in the examples.

A: Which figure are you suggesting to add?

Overall, the paper is missing some discussion and related work, but represents a significant improvement for generating solutions for QDec-POMDPs.

Reviewer 2:

The paper presents an alternative planning approach for QDec-POMDPs (which are a deterministic analogue of the standard Dec-POMDP model). To reduce the complexity of planning, the authors present an iterative approach (which is not complete as it may fail to find valid plans) via reduction of the given multiagent planning problem to a single agent contingent planning problem. The authors also present a number of heuristic techniques that can enable agents doing such single-agent planning to converge to a valid multiagent plan.

The work is relevant for IJCAI. The paper is generally readable. Solving QDec-POMDPs is indeed very challenging, so the motivation to come up with better and scalable approaches is also good.

Some of the main areas where more clarifications are needed are discussed below.

First the authors should clearly identify the class of problems where the current approach is expected to work. A major highlight of the proposed approach is that it is not guaranteed to find valid plans even if they exist. So it necessitates a clear class of relevant and significant problems where the current approach is expected to work. A detailed discussion regarding this point is currently missing in the introduction.

A: We agree with this comment. We believe that in Ergodic domains (no deadends) our approach is complete, but we do not yet have proofs, and hence, are hesitant to make such claims.

 The writing quality (wrt the organization of the content) needs improvement at several places. The paper has several moving parts in the overall approach. However, there is no systematic point-by-point description of all the techniques/heuristics the authors have to employ.

A: We are not sure that we understand the suggested improvements here.

As a result the overall picture regarding the complete approach, its runtime/memory complexity is very hard to judge.

Some examples are below:

-       In section 2, it is better to describe the model using bullet points highlighting various aspects.

-       Section 2 contains both model description and policy description/execution. Certain crucial things in the policy execution are unclear. E.g., during executing a policy, agents maintain a belief. Can the set of possible states in this belief be exponential? What is the complexity of maintaining and updating this belief? What is the precise procedure to update this belief? All such things needs a precise description, in addition to the procedural way in which they are currently described.

A: There is no need to maintain a belief during policy execution, when policies are represented as trees (or graphs, or automata). We maintain beliefs during planning using regression, but this is a part of the contingent solver.

-       Several subsections of section 3 need a more clear description. Section 3.2 describes a number of additional heuristics to extract a collaborative plan using techniques such as addition of time to the plan. Please include such additional details into algorithm 1 that gives a better overall picture.

 A: We had to keep our pseudo code extremely short sue to the very limited space. We agree that much is missing.

Empirically, a discussion and comparison against a recent related approach for QDec-POMDP is missing:

Knowledge-Based Policies for Qualitative Decentralized POMDPs. AAAI 2018.

A: This is a very interesting paper that suggests a new way to model policies for QDec, but does not suggest an algorithm for finding such policies. Hence, no comparison can be made. On the other hand, this is an excellent example on how papers on a relatively unexplored problems, such as QDec, do not adhere to the standards in more explored domains, such as Dec-POMDPs. It is difficult to see a paper suggesting a new policy representation for Dec-POMDPs with no algorithm accepted to AAAI.

It is better to compare the current approach to this directly related work rather than comparing against Dec-POMDP solvers which are solving a different problem. Furthermore, the authors should state clearly what is the advantage of their approach over this relevant previous work.

A: Again, the only previous algorithm is the compilation approach from the original QDec paper, which we compare against.

As the authors also acknowledge, the approach is not guaranteed to find valid plans (the approach is incomplete). This is a major discussion point. It is important to discuss domains (and their particular characteristics) where the current approach would work. The authors state that such a discussion is left for future work, which seems an unsatisfactory answer because if practitioners do not know when and where the approach is going to fail/work, it makes it very difficult to judge the significance of the work.

A: IMHO, incomplete algorithms for very difficult problems are much better than complete algorithms that cannot solve reasonably sized problems. We believe that this approach was also embraced by the Dec-POMDP community when preferring to investigate “approximate” Dec-POMDP solvers.

Reviewer 3:

This paper addresses cooperative multi-agent planning problems under imperfect information. A stochastic setting that previously deals with such a goal is the decentralized control of partially observable Markov decision process (Dec-POMDP). Unfortunately, this framework does not allow us to model its qualitative alternative. In other words, what if we use possibilities rather than probabilities to formalize the non-determinism of Dec-POMDP? This question gives rise to Qualitative Dec-POMDPs (QDec-POMDPs), a framework, Brafman et al. introduced recently. It turns out that both Dec-POMDPs and QDec-POMDPs are NEXP-Complete in the worst case. To tackle this complexity barrier, the authors suggest yet another variant called deterministic QDec-POMDPs. They further introduce an iterative approach to compute possible solutions to QDec-POMDPs.  The authors run experiments on modifications of the box-pushing benchmark, which fits assumptions made in deterministic QDec-POMDPs.

Solving cooperative multi-agent planning problems is a significant topic at IJCAI. This problem is of interest to researchers in classical planning, constraint programming and sequential decision making. These recent years have seen some tentative to build a bridge between those approaches, especially within new frameworks such as QDec-POMDPs. However, it is unclear to me what do deterministic QDec-POMDPs formalize that existing models in deterministic settings, such as those in planning or constraint programming, could not? A careful look at the assumptions made about deterministic QDec-POMDPs reveals that we are close to standard multi-agent planning problems rather than QDec-POMDPs.

A: We strongly object to this comment. Deterministic QDec maintains the major difficulty of having to consider the joint observations of agents. That is, agent 1 must consider all possible observations that agent 2 has seen. This is the major difference between fully observable MA planning, or single agent partially observable planning, and QDec.

The main assumption of deterministic QDec-POMDPs on the top of QDec-POMDPs is the fact that both transition and observation dynamics are predictable. With this assumption as a background, it seems to me that the problem restricts to a classic multi-agent planning problems. It seems to me that the partial information that is key to (Q)Dec-POMDPs no longer holds. In (Q)Dec-POMDPs, an agent acts without any information about what the other agent sees or plans to do.  In deterministic QDec-POMDPs, the sequence of actions is the total information necessary to act optimally for the team. In other words, under the knowledge of the sequence of actions, the information contains in the observations is unnecessary, mainly because the process is deterministic, one can predict those observations.

A: this is not true, the observations reflect the unknown initial state, not just the actions.

Assuming the previous argument holds, then deterministic QDec-POMDPs are equivalent to standard multi-agent path-finding.

A: QDec are very different than MA path finding (e.g. the work of Sharon, Stern, and Felner), where everything is fully observable, and there is no uncertainty concerning the initial state, and hence control is typically centralized.

As a consequence, the contributions of the paper regarding the representation of policies as trees is useless. Perhaps a quick attempt to analyze the complexity of deterministic QDec-POMDPs can help understanding what makes this problem different or close to existing ones.

Suppose for a moment that deterministic QDec-POMDPs are useful for some applications. A reader may wonder why dynamic constraint programming methods fail to address deterministic  QDec-POMDPs? The authors adapted a standard algorithm---called Joint Equilibrium Search Policy (JESP)---in Dec-POMDPs to deterministic QDec-POMDPs, which is fine. A similar idea was recently proposed for multi-agent planning problems in Scaling-Up Multiagent Planning: A best-response Approach (ICAPS'11 by Jonsson and Rovatsos). So the algorithmic scheme of IMAP (the proposed algorithm) is not new, and a reader may wonder why existing approaches such as that of Jonsson and Rovatsos do not apply in deterministic QDec-POMDPs?

A: Jonsson and Rovatsos work on fully observable MAP, which is indeed very similar to MA path finding, but very different than QDec, where the distributed partial observability is the key source difficulty.

Regarding the experiments run for the paper, the only benchmarks used are derivative of one and only one standard benchmark from Dec-POMDPs.

A: This is not true – we also adapted an MA planning problem, Rovers, to QDec, and report results on this benchmark as well.

Once again, it seems to me that this is just another classical planning task performed with multiple agents, and there is no reason to add yet another framework. The authors compare IMAP to standard solvers for Dec-POMDPs. This comparison is not fair at all. The authors compare probabilistic approaches with a qualitative one on  deterministic benchmarks. It would have been more informative to take standard classical planning/constraint programming approaches to demonstrate their limitations.

Reviewer 4:

Relevance: 8

The paper is clearly in the scope of the conference.

Significance: 5

The algorithm introduced for solving QDEC-POMDPs seems to perform well on two case studies compared to the standard algorithms of Dec-POMDPs. However, it is rather difficult to compare it in general with existing algorithms (as the authors suggest on page 5) and there is neither a theoretical result regarding this algorithmic proposal nor a comparison with related works in the game theory literature, (specially the results about the undecidability of multi-player games, see below).

Originality: 5

The algorithmic proposal is interesting, but the design choices of the solution seems are not really motivated, even if they make sense. For example, nothing is said about the ordering of the agents in the overall algorithm, although this might impact the efficiency of the algorithm depending on the goals that each of them can achieve.

A: We agree that better ordering may result in computing solutions faster, and we are considering such extensions. We cannot fit everything into a 6 page paper.

Technical quality: 6

There are few formal definitions, only one high-level algorithm and no result, so it is difficult to assess the technical quality of the paper. However, one can spot some inaccuracies:

- the vector of individual action sets is not defined: \vec{A}:=(A\_i,…, A\_n).

- the inclusion at the bottom of page 2 shouldn’t be rather a difference ?

- G\_i^T(g) is not formally defined (it corresponds in fact to the t\_g of Section 3.4)

A: we thank you for spotting these issues.

Clarity and quality of writing: 7

The paper is well written, although there area number of typos (like at the beginning of Section 3.1).

Scholarship: 4

A: we do not focus here on complexity results. We do not expect to have different results than the original QDec paper.

The paper is not really well-situated with respect to the game-theoretical literature, the verification literature and the supervisory control theory literature. There are many results about undecidability of multi-player imperfect information games that are not recalled, a basic reference being:

- Reif J.H. (1984), The complexity of two-player games of incomplete information. Journal of computer and system sciences, vol. 29(2): pp. 274–301.

A: Games are in general non-cooperative. QDec are fully cooperative, i.e. rewards are joint.

The problem the authors address is an offline distributed planning problem without communication. This work of Dima from the verification literature is also relevant:

- Dima & al. (2010), Model-Checking an Alternating-time Temporal Logic with Knowledge, Imperfect Information, Perfect Recall and Communicating Coalitions.

A: There are obviously links between verification and planning under partial observability, and intuitions may be drawn from one field to another, but using algorithms from verification for QDec is far from being straight forward, and it is unclear, without a thorough investigation, whether these ideas would work well.

There are also lots of work in the supervisory control theory regarding « decentralized control»  that addresses similar problems, a basic reference being:

- Ramage and Wonham (1989), The control of discrete event systems, proceedings of the IEEE, vol. 77, No. 1.

A: Indeed, there are related topics in planning and OR. These are not sufficiently explored in all planning literature. It is unclear to us why the specific referenced paper, dealing with systems with discrete events, where there is centralized control through disabling events is deemed particularly relevant to our line of work, except in the very broad terms.

Overall score: 5

As said above, there is no theoretical result, the comparison to related work is rather shallow and the algorithm proposed seems rather « ad hoc » to me. Also, the underlying assumptions regarding the planning setting and the constraints implicitly imposed by the algorithm on the plans produced are not explicitly given.

A: This work suggests a practical algorithm for the Dec-POMDP framework that was previously suggested and analyzed theoretically (2013). As such, it is an empirical paper, not a theoretical one.