TODOs for GBG-Framework – WK/05/2017

## From TR-GBG.pdf (is commented out there)

* The n-tuple agent developed for C4 (Connect Four) needs to be ported to GBG.
* Allow only trained agents to be saved.
* Optional game visualization and game logging during competitions as well
* Clarify: Is the parameter data flow safe, if we issue a ’play’ or ’compete’ for 2 agents of same type but with different parameters?

## Solved items

* OK Bug fix: The current LoadAgent version ‘hangs’ if an agent with incompatible serialID is read (TTT agents TDS & TD-NTuple).
* OK Bug fix in TDAgent.java: We needed to replace the hard-coded   
  private int inpSize[] = { 6, 6, 10, 19, 13, 19, 0, 0, 0, 9 };  
  by the flexible m\_feature.getInputSize(m\_feature.getFeatmode()).
* OK TD-pars tab: Constrain ‘Feature set’ (= value for featmode) to the set of allowed featmode values for the current game (m\_feature.getAvailFeatmode()) and set the initial value to a sensible value 🡪 KG fixed this and generated a pull request
* OK: BUG1 + BUG2 fix for MCAgent: In some versions of Hex, the MC agent does not choose the best action, but **instead the worst** action (!!) This happens for (3x3, 1.), (4x4, 2.), (5,5, 1.), … and so on. The same behavior for TTT (3x3, 1.) – This was due to 2 errors: (1) in RandomSearch there could be actions on newSob even if newSob is already a game-over-state (BUG1) and (2) newSob.getGameScore has to be called with the referingState sob to get the sign right in every case (BUG2)
* OK Improved TR-GBG around game score and game value: Made it simpler!
  + Concentrate Chap. 3.3 on so.getGameScore() and so.getGameScore(sob). Explain the difference in pictures. Give examples showing explicit values for both functions.
  + Move pa.getGameScore(so) to Chap. 3.4 (Game Value), because it is a game value, not a game score (think about renaming it to pa.getGameValue).
  + Move the mind-buggling complicated so.getGameValue() and pa.estimateGameValue(so) to appendix.
* OK XArenaFuncs::multiCompete writes Arena.comp.csv now to agents/<gameName> (and not to root of GBG)
* OK XArenaFuncs::multiCompete has the evaluator mode hard-wired (9). This might fail for other games where the evaluator mode 9 is not present. Extend it that it fetches static Evaluator.getDefaultEvalMode().
* OK Why has XArenaFuncs::multiCompete the evaluator part excluded in case of agent MCTS? – This was due to a wrong (incomplete) implementation of MCTSAgentT::getScore (the part with nextActionScore, same as in MCAgent, was missing). -- Additionally, there was another bug in both MCAgent and MCTSAgentT: Both would have problems when getScore(sob) was called with an already game-over sob (as it regularly happens inside Evaluator9). Then MCTSAgentT resulted in a NullpointerException, since the tree of a root state with game-over condition is never expanded in treePolicy. The fix is to check in getScore(sob) for sob.isGameOver(). If true, return directly sob.getGameScore(sob). Now both bugs are fixed and MCTS is correctly evaluated in multiCompete.
* OK Improve Evaluator chapter in TR-GBG.
* OK Write for TR-GBG.pdf an appendix on N-tuples (with figures): Explain for what the several functions in XNTupleFuncs are needed.
* OK Types.ACTIONS: replace enum (which needs to be extended whenever we need more ACTIONS) by some class construct
  + a class having an int as member
  + we delete fromString, fromVector, they are never used in GBG
  + older logs (with enum ACTIONS) are not readable anymore
* OK Bugs in Hex (for KG):
  + OK When playing a human-agent game, then the scores of the agent are not shown during play (they are however shown when replaying it with the game log)
  + OK Hex: extend to logging with subdirs 🡪 KG
  + OK Under Hex, both Arena and ArenaTrain open on “Load Agent” the wrong directory (it should be agents/Hex/<num> where <num>=BOARD\_SIZE)
* OK Added   
   static final long **serial****VersionUID**

to serializable classes to ease the serialization process. Then the serialVersionUID is no longer compiler-dependent

* (OK) MCTSAgent-Design-Flaw: It has params, SinglePlayer m\_mcPar and the single elements 🡪 too much!!! See if we can delete params 🡪 yes, we did. And we let the setters for single params change m\_mcPar as well. And we added getters for MCTSAgentT. Do we need setters in MCTSAgentT (or do we go always through constructor)?
* OK MCTSParams: bug in getK\_UCT(): intValue() 🡪 now corrected to doubleValue()
* OK MCTSParams tab: when changing params and directly saving MCTS, the params are not taken. Only after a “Play” the params are taken over (!) 🡪 saveAgent in XArenaMenu now uses fetchAgents first!

## Open items

* Prepare TR-GBG.pdf for publication on CIplus server.
* Saving (serializing) of things (logs, agents) is still not safe, if something changes in the classes being serialized. Is it possible (with minimal effort) to read older versions of serialized objects as well? – It has become better with the proper definition of [serialVersionUID](#serialVersionUID). A further improvement might be to catch an InvalidSerialVersionUID exception and react on this with transforming from an older version?
* During TDNTupleAgt-training: Replace the Minimax-evaluation by proper general evaluation (replace JFreeChart plot title)
* Should there be game-specific functions
  + - BoardState toBoardState(ACTIONS) and
    - ACTIONS fromBoardState(BoardState

where BoardState is a game-specific class (e.g. holding row and column int for a game like Hex)?

* (OK) Think about the game logic for so.getGameScore(referingState) in 3-player games 🡪 see [notes\_MCTS.docx](file:///C:\WUTemp\FH-MassenDaten\svnSoma\trunk\doc\CaseStudies.d\201314.d\CIG2014\MCTS.literature\notes_MCTS.docx), Chap. “Getting The Score For n-Player Games”. – Still some thinking about intermediate game scores to do.
* If MCTS has several actions with the same value (all are a ‘Win’), then it does not take the shortest path to victory. Instead it takes a random among all winning moves. 🡪 KG will think about an add-on to value function which breaks ties in favor of a shorter path.
* An Evaluator object is currently constructed in several places (multiCompete, multiTrain, train, and taskState TRAIN, menu item ‘Quick Evaluation’ (fct evaluate)), often with different objects (different evaluator modes), and with modes tied to game TicTacToe. Make the modes as simple as possible, then generalize it in such a way that it is for arbitrary games. Perhaps with user-adjustable mode selectors in the ‘Other pars’ tab, perhaps with sensible defaults set in the files defining constants. 🡪 see [#Evaluator concept](#_Evaluator_concept)
* Some elements of measurement in methods train, multiTrain, and compete are still TicTacToe specific or at least 2-player specific. Generalize them to arbitrary 2-player games and later to arbitrary n-player games.
* Update the Help text for the various GUI elements
* Improve GUI layout for Arena and ArenaTrain
  + Arena: fill the empty space with s.th. sensible (e.g. game name)
  + Arena: announce which column is for which player (‘Param X’ and so on from ArenaTrain is missing here). Optional: show the color of the player
  + Move button MultiTrain in the ‘Agents trained’ row and move button ‘Logs’ in the blue buttons row.
  + Arena: skip the two rows ‘Train games’ and ‘Agents trained’
  + Arena: there is no way to select Human Player via agent select box. Should we allow to select all agents (or all agents which need no training) via select box?
  + think about nicer colors
* Bugs in Hex (for KG):
  + When replaying a human-agent game with the game log, then any ADVANCE will move two plies foreward (one agent, one human ply). When replaying an agent-agent game, it is correctly only one ply per ADVANCE.
  + If agents/Hex does not exist (Arena.comp.csv), then create it
  + Multi-Competition: seems that wrong agents compete (MCTS when MC is desired) 🡪 check multiCompete, check whether its training is o.k.
* Things for MCTSExpectimax (for JK):
  + OK TreeDepth <= 3 leads to NullPointerException (s.th. around   
    return uct().treePolicy().treePolicy();) 🡪 Bug is gone in new version
  + Why is TreeNode-constructor called exactly nIterations (3500) times (at least for TreeDepth=10)?
  + Why is MCTS(Expectimax) with TreeDepth=1 not as good as MC? Which role plays K (UCT)? Does it help to set it to large values?
* ? Competitor class to make the code parts dealing with competitions more similar / more safe
* TR-GBG.pdf: new sections on important reoccurring questions
  + How to find good features?
  + How to evaluate an agent?

## equals() or hasEqualState()?

* Design discussion: Is it safe to have StateObservation::equals() for checking that two states are the same?
  + pro: If we have it this way, we can use all other code (e.g. Set, List) which builds upon equals()
  + con: There is always the default implementation Object::equals() which has a different meaning (two different objects can have the same state). A developer of a new class implementing StateObservation might forget to re-implement equals()
* Another possibility would be to add a new function   
   hasEqualState(StateObservation arg0)

to the StateObservation interface

## Evaluator concept

Where and how is Evaluator used?

* ArenaTrain, derived task TRAIN: after training, an Evaluator with mode=2 is constructed and evaluated 🡪 should be mode getQuickEvalMode()
* XArenaFuncs, train() & multiTrain():
  + m\_evaluator1: Evaluator with mode 9
  + m\_evaluator2: Evaluator with mode 2
  + m\_evaluator3: Evaluator with mode 0 (only multiTrain())
* XArenaFuncs, multiCompete():
  + m\_evaluator1: Evaluator with default mode (for X)
  + m\_evaluatorO: E valuator with default mode (for O)
* XArenaMenu, evaluate(), which is only called by ‘Quick Evaluation’:
  + evaluator2: Evaluator with mode 2
* Disadvantage: The mode values are hard-coded (perhaps only appropriate for TicTacToe)
* The different compete menu items should be disabled for 1-player games and generalized appropriately for (n>2)- player games
* Arena2048, ArenaTrain2048: both have some strange makeEvaluator functions which compare the constant (!) ConfigEvaluator.Evaluator with 0 or 1 (🡪 compiler warning & can only be changed after recompile)
* Nicer design: it would be better to use parameter mode of Evaluator constructor / of makeEvaluator()
* EvaluatorHex: eval\_Agent chooses between three different functions competeAgainstMinimax, competeAgainstMCTS, and competeAgainstRandom at compile time
* Nicer design: Have three different Evaluator objects which decide via a mode variable which competeAgainst… function is used. This has two advantages
  + Several competeAgainst… functions can be evaluated in parallel
  + The decision, which Evaluator object should be taken can be made at run time

A more general Evaluator concept:   
How should the new Evaluator usage be?

* There are five new helper functions in Evaluator interface
  + int getQuickEvalMode()
  + int getTrainEvalMode()
  + int getMultiTrainEvalMode()
  + String getPrintTitle() (e.g. "Avg. success rate (randomAgent, best is 0.9): ")
  + String getPlotTitle() (e.g. "success against Minimax")
* When starting ‘Quick Evaluation’, only an Evaluator with mode=getQuickEvalMode() is constructed and evaluated
* When starting TRAIN, there might be two evaluators, if getQuickEvalMode() and getTrainEvalMode() return different int’s. If they return the same int, only one evaluation is done. In any case, the QuickEval Evaluator is used to add a point to the JFreeChart plot and to check whether goal is reached (and if so, break out of training loop). The other evaluator with mode getTrainEvalMode() is only used for additional reporting at the end of training.
* When starting MULTITRAIN, there might be similarly one, two or three evaluators, depending on whether getQuickEvalMode(), getTrainEvalMode() and getMultiTrainEvalMode() return different int values.
* The helper functions getPrintTitle() and getPlotTitle() are for convenience, to customize printout and plot axis title.

So the simplest usage of the Evaluator concept would be: have only one mode value implanted and let getQuickEvalMode(), getTrainEvalMode() and getMultiTrainEvalMode() return this one mode value.

A possible add-on: Have in the “Other Pars” tab a selector “Quick Eval Mode”, with initial value getQuickEvalMode() and with possibility to change it to any available eval mode (what getAvailableModes() returns). With the effect that the mode value used in ‘Quick Evaluation’, TRAIN and MULTITRAIN changes accordingly.

A possible later add-on: Let eval() return an object of class EvalResult which has getters getSuccess(),goalReached(), getLastResult(), getMsg(). Why? Consistent object, information is kept together.