Paper Review form

30Oct2014 - 2nd Review...

Only suggestion :
p14h0.5 References - usually the section title “References” appears before the listing

+-----+
Neural Networks Paper Reviews
http://ees.elsevier.com/neunet/

Paper ID : Paper ID# and Authors' names removed for privacy
Title : Fully probabilistic control.odt
Assigned : ????
Due : 16Sep2014

For each question, please use the following scale to answer (place an X in the space provided):

RATINGS

1 Superior
2 Good
3 Fair
4 poor
5 Not applicable

Quality of Methodology : 1
Quality of Work : 1
Soundness of Conclusions : 2
Significance of Subject : 1
Clarity : 1
Organization : 1
Priority Rating for Publishing in Neural Networks ("1" is highest) : 1

Is the abstract, and are the figures, legends, and references acceptable? If not please explain:

Yes - although the key part was the clear and methodical way that the mathematical expressions were developed, and the basis for them.

Please provide a brief and compelling argument supporting (a) your recommendations and (b) the above ratings:

See my comments to the authors.

****

This reviewer's personal approach, nomenclature examples:
p1c1h0.8 = means page 1, column 1 80% of the way down the page (very approximately)
C2. = means Comment section #2 WEAKNESSES (note that actions by the authors are NOT required for the points)

++---------------------------++
ACTIONS REQUESTED OF THE AUTHORS

I have NO critical changes to request of the author. In my opinion it is a very strong paper with important concept developments, and it is ready for acceptance.

In the “C.4 Details” section, there are some unimportant s(mostly stylistic) suggestions which the author might consider or not to use.

SPECIAL Note : I did NOT do an initial review for this paper, which is well-written and sound. Furthermore, the author appears to have well addressed the points raised by the original reviewers. Because of this, it is unfair to the author to add additional review details at this stage, and there should be NO necessity for the author to respond to my points below, which would delay publication.

Standard NOTE: The points above are the ONLY points that I request that the authors address. From here to the end of the review, there is no requirement for the authors to make any changes to the paper, nor is there a need to respond to me about those points. I provide comments that they may consider at their own discretion, and it is NOT my intention that any of the points below have to be addressed (other than those listed above). I am very afraid that authors feel that they are obliged to answer or make changes, which could waste far too much of their time, and several of my comments may be speculative and may not even be correct!

****
****
COMMENTS ONLY - actions by the authors are NOT required for the points listed below, to the end of the review. Perhaps some of these comments will be helpful in some way.
(Main paper contributions, positive aspects, observed deficiencies, and suggestions on how to improve them:)

PERTINENCE of the paper for the “Neural Networks” journal :
As per the Aims and Scope” listed on the copyright page of each issue of the journal :
“... Neural Networks welcomes high quality submissions that contribute to the full range of neural networks research, from behavioral and brain modeling, learning algorithms, through mathematical and computational analysis, to engineering and technological applications of systems that significantly use neural network concepts and techniques. ...”

This paper is well-suited for NN Journal.

++---------------------------++
C1. STRENGTHS OF THE PAPER:

Overall, this is VERY interesting and is work well done. The paper is clear, well organised, and the mathematical developments are solid.

Furthermore, the results seem to me to be very important, at the very least in a theoretical sense, and in a very practical sense useful at the least for difficult systems where [real-time adaptation, learning, evolution] isn't required. Perhaps even that is doable at a later stage, but I can't see that now, unless abductive (pattern matching, metaphors) are used?

Collectively - throughout the paper the author provides strong arguments for Fully Probabilistic Design (FPD) control :
p1h0.6 "... Robust control strategies in the absence of reliable world models and in the presence of strong uncertainty must be designed for practical control to take place. ..."
p3h0.4 "... Nevertheless, using single model approaches for estimating the involved pdfs is restrictive in many real world applications that are characterised by strong nonlinearity, multimodality, and uncertainty. ..."
p3h0.55 "... In this framework, a weighted sum of local models is used to model the non-linear dynamics of the system. In effect, every local model is a representation of one particular operation mode. ..."
p3h0.8 "... The method proposed in [15], however, is constrained by its high computational efficiency. ..."
p4h0.6 "... The proposed method describes a general way of how to understand and incorporate uncertainty in deriving optimal control strategies. It emphasizes that when the density function parameters are dependent on the input values are dependent on the input values, not only [SHOULD] the expected value of the Kullback-Liebler distance be minimised, but also the variance of its cost function. ..."
p5h0.1 "... To emphasize, generalised probabilistic controllers proposed in this article provide a pragmatic method for effectively and robustly controlling complex stochastic dynamical systems under uncertain conditions. They take knowledge of uncertainty into consideration in the derivation of the optimal control law. This represents the novelty of the new generalised probabilistic controller proposed in this paper. ..."

p18h0.75 paragraph starting "... The system state and desired state ..."
p19h0.0 Figure 2 (this figure is very effective!)
The author's Simulation Example clearly shows advantages of his GenProb-DHP over the "standard Probabilistic DHP" adaptive critics

Collectively - throughout the paper the author provides strong arguments for a Mixture of Density Networks (MDNs) :
p4h0.2 "... Mixture density network from the neural network field [4], [11] will be used in this paper to estimate the required complete pdfs, as multiple components of Gaussians, such that their parameters are state and control dependent. ..."
p9h0.3 "... from the MDN. A method of interest in multimodal control applications is the most probable component from the MDN corresponding to the most probable branch. ..."

plus other comments in the paper...

The need for the adaptive critic approach, and it's smooth blending into the FPD, is well explained and done. For example :
p16h0.1 "... critic network and system state density models become available. The implementation of this two stage optimisation method can be performed efficiently by utilising the modular approach constituting of functional modules and algorithmic modules [8], [9], [18]. The main functional modules are the action and critic networks. Algorithmic modules on the other hand include the computation of the desired critic value (29), computation of the optimal control law (34), and the update of the networks parameters. Each of these modules can be modified independently from other modules, thus facilitate fast and reliable implementation. ..."
That was nicely stated, and is a great point.

p16h0.33 Nice description of the overall GenProb-DHP FPD process (steps 1 through 10 with iterations). I do make the comment in "C.4 Details" that the description for the simulation example could follow the same numbered breakdown (even though it's an obvious match).

++---------------------------++
C2. WEAKNESSES: (again, changes to the paper are not require for these comments)

The following comments aren't really weaknesses, as they go beyond the scope of the paper, and are not essential to its core theme. These are more like "items that would be nice to see" in future papers, and they are certainly NOT issues that should delay the publication or require more work of the author.

p19h0.0 Figure 2 - As mentioned under "C.1 Strengths", the author does a good job of demonstrating the advantages of the FPD approach (GenProb-DHP) over the "standard fully Probabilistic DHP adaptive critic approach [Herzallah & Karnay 2011]". However, there are several questions about the results :
- What is the comparison between errors in the "stabilised plateaus" of Figures 2b and 2d? By eye, the FPD is superior, but how does it compare quantitatively, and how close is it to some kind of measure for "ideal performance" (which would take into account the overall system uncertainties - perhaps tricky to define as a measure?)?
- What would the result be for a normal non-Probabilistic GDP approach? Here I assume that the example was selected may not be amenable to the normal approach, and this may be available in previous papers (maybe [Herzallah & Karnay 2011] [Smidl etal 2005]). But even if so, it would be interesting to see graphically in Figure 2.

I suspect that the author has seriously understated the advantages of his FPD GenProp-DHP, going beyond he many points made (see section "C.1 Strengths" above), even though that may be implicit and obvious to experts in the domain.
- ongoing indication of uncertainty - you don't mention it in the paper, but a BIG problem with many techniques is that one is not given a solid indication of the uncertainty of the [prediction, control, pattern match, etc]. Your approach could give much more braod and deep measures of this, allowing interpretation of the KINDS of risk the system faces.

- "robust" versus [robust, reliable, survivable, risk-reduction] - I feel that sometimes it is important to distinguish between these concepts.
- Linguistic indicators for modes - Can the resulting FPD provide a basis for an [understandable, automatic] linguistic description of the FPD status? Can this help "operators" better understand the system?
- Multiple Conflicting Hypothesis – it seems that GenProp-DHP could easily be adapted to entertain multiple hypothesis and flip quickly between them when divergences with the currently selected hypothesis become problematic. Moreover, where multiple hypothesis can work well, one might select the hypothesis, not with the maximum expected benefit or lowest maximum penalty, but on the basis of one that leaves the most options open for later.

Points like these (even if my comments above are erroneous) may be important to point out to readers in future publications on this topic?

++---------------------------++
C3. QUESTIONS: (no need to answer)

p8h0.85 "... Finally the ideal joint pdf I h(xt , ut | x t−1) = I s(xt | xt−1 , ut) I c(ut | x t−1) is assumed to be independent of the component kernels of the actual joint pdf h(xt , ut | xt−1) = s(xt | xt−1 , ut)c(ut | xt−1). For example if a tracking control problem is considered, a reasonable form for the ideal state pdf, I s(xt | xt−1 , ut) = If(xt | ut , xt−1) would be a Gaussian pdf with its mean value being set to a desired value and variance to the non–reducible variance of innovations of the active component. It specifies where the system state will be of high probability. Similarly, the ideal control input pdf I c(ut | xt−1) = Ig(ut | xt−1) can also be assumed to be Gaussian which expresses where the system control input should be. ..."
>> Reviewer :
1. This seems to be a key assumption that is somewhat counter to the basic theme of the paper of FPD, but which renders the subsequent analysis tractable. I don;t have any feel for how good this assumption may be for the difficult systems for which FPD is intended.
2. Perhaps the author has back-analysed the results of the Simulation Example to "see" the actual joint distributions, and has some insight as to when this assumption could lead to trouble?
3. To me, it is LIKELY that the s and c distributions are STRONGLY inter-dependent on some occasions/ situations, leading to strongly non-Gaussian distributions (including multi-modal, Chi-squared-like, etc). This could easily occur even if, over all of the problem space, the overall joint distribution is Gaussian.
4. The MDNs are NOT used for the joint distribution - will this cause trapping in local minima, and serious "gaps" between realistic data and the idealistic Gaussian assumption - gaps which may tend to cumulate over time (rather than cancel like noise) with respect to the optimal solution path?

p14h0.0 critic, p15h0.0 controls - Taylor series truncation to second order :
This was a nice approach. Where might this cause problems? It doesn't appear that going with more terms would be an impediment to solving for the authors' FPD, although it would make it more complicated and solution would take more time (not necessarily an unacceptable amount of time?). But I am guessing here.

Equation (35) may not test the second order truncation to any great extent?

Non-stationary systems with [state, phase] changes, or heteroscedastic nature :
- Presumably the FPD can account for this because of the way that the approach is defined, but that isn't entirely clear, nor how far that might be pushed and still have robust control (i.e. would a more general form of FPD be required for strong multi-[phase, state] systems? Clearly the [states, phases] must exist in the data.
- Feldkamp/ Prokhorov showed years ago the ability of RNNs to [identify, accommodate, control] state/phase changes, which must also be the case here?
- Even if the system runs into a problem with this, the FPD GenProp-DHP approach will probably provide the information to warn that it has happened, or about it's potential.

Target pathways - How limited is this approach to fixed paths, as opposed to curved trajectories or dynamically-changing targets? It seems that it could be quite adaptable within limits, as long as pre-batch-training covers the changes.

Real pdfs - what do they look like for [critic, control, overall GenProp-DHP FPD]?
p14h0.8 "... The generalised probabilistic DHP adaptive critic method computes the optimal control inputs by deriving Equation (25) with respect to the control inputs and setting the derivative equal to zero. This results in the following optimality equation for computing the generalised probabilistic control values: ..."
- Will this basic assumption of p14h0.8 result in the controls being "stuck in local minima" for multi-modal real pdf distributions, thereby losing effectiveness? If so, it should be easy to check, after training, if that might be a problem?
- Even if this is an initial problem, how can one overcome it?

problems of dimensionality - how does this scale?

correntropy - Jose Principe's concept. Although this does not relate directly to this paper, it came to mind when reading through parts.

p18h0.75 How many "outer" (steps 4 to 10) iterations were required for solution, and for a sampling of start, middle, end] outer loops, how many "inner" (critic, control) iterations were required? Actually, it may be easier to measure run times - which brings up the question of time to solve for the StanProb-DHP versus GenProb-DHP.

++---------------------------++
C4. DETAILS and GRAMMAR: (again, changes to the paper are not required for these suggestions)

The paper is well written and is easily understood, so I have few suggestions to make. These tend to be stylistic, and may not be better than the authors' original version, so they should only be considered if to the authors' liking.

p1h0.33 change to "... Using the proposed "Generalised Probabilistic Dual Heuristic

Programming" (GP-DHP) controller, ..."
p3h0.
I just thought it would help over the long term to have a full acronym for your approach, but that's up to you. "GenProb-DHP" is another possibility among many others, with the advantage that it gives a better hint of what you are doing, but it's a bit long, and scientists may not like the slang form (I like it better, as it reminds me of what you do, but then, I have no taste).

p1h0.9 Citations : as a general comment, NN journal requires the format "... [Astrom 1970], [Werbos 1982] ..."
This applies to all citations.

p2h0.75 change to "... Here -ln(gamma(x_t-1)) ..."

p3h0.55 change to "... Depending on the problem being handled, ..."

p3h0.8 perhaps change to "... recurrence equation analogy of ..."
But my suggestion does not seem correct. Perhaps "... equivalent to ..." or "... implementation of ..."?

p3h0.85 change to "... this necessitates sequential evaluations and storage of the expected ..."

p3h0.9 change to "... even more arduous for the majority of practical systems, which are non-linear and non-Gaussian in nature. Most importantly, ..."

p3h1.0 change to "... to be input and state independent, limiting the resulting control strategies to deterministic, certainty-equivalent systems. ..."

p4h0.2 change to "... Mixture Density Networks (MDNs) ..."
It doesn't hurt to repeat the acronym here, even if clearly stated in the Abstract.

p4h0.25 perhaps change to "... required complete pdfs as combinations of Gaussian components. ..." perhaps in parenthesis "(like radial basis functions (RBFs))"?

p4h0.5 change to "... not only should the expected value of the Kullback-Liebler distance be minimised, but also ..."

p4h0.7 change to "... such that the system and control models' uncertainties are taken ..."

p5h1.0 change to "... the conditional pdf of the state, xt, should be modelled based on its previous state, xt-1, and the control input values, ut. This paper applies MDNs to model the general pdf of the system state from process data. ..."

p6h0.1 change to "... the multiple models approach which will be discussed ..."

p7h0.2 change to "... and applied control inputs, are then applied ..."

p7h0.15 change to "... Measured process data, states and applied control inputs, are then used to train ...”

p7h1.0 change to "... the pdf of the system states, a paired control input kernel function must be designed [Herzallah 2012]. ..."

p9h0.1 change to "... definitions of the state pdf given in Equation (6), the control input pdf given in Equation (9), and the assumed Gaussian form of the joint pdf between them, an approximate ..."

p9h0.2 change to "... The FPD control solution depends on the method which is selected for calculating the ..."

p9h0.25 change to "... literature [Bishop 1995], [Herzallah & Karnay 2011], which can be generalised ..."

p9h0.6 change to "... control input pdfs as given by Equations (11) and (12), yields the optimal ..."

p12h0.25 change to "... as the Generalized Probabilistic Dual Heuristic Programming (GP-DHP) adaptive critic method. ..."

p13h0.6 maybe change to "... Analogous to the standard fully Probabilistic DHP adaptive critic approach [Herzallah & Karnay 2011], a critic network ..."

p13h0.95 change to "... this equation cannot be obtained analytically due to ..."

p15h0.9 change to "... is a two stage process. ..."

p16h0.3 change to "... the update of the networks' parameters ..."

p16h0.33 change to "... thus facilitating fast and reliable ..."

p16h0.6 possibly change to (point 5) "... Specify the idealised Gaussian joint density function between the state and control inputs. ..."
I may be misinterpreting or misunderstanding this, but the control and the critic distributions use Gaussian (RBFs?) to estimate real distributions, not to come up with idealized Gaussian pdfs. It is the JOINT pdf which is idealized as a Gaussian?

<<-----+
On page 17, I suggest emphasizing your excellent 10-step process description on page 16, by separating the Simulation Example into numbered sections :

p17h0.5 change to "... Probabilistic control input. Following the steps in Section IV, page 16 above : ..."

p17h0.5 change to "... 1. Estimate the general pdf ...
As a first step ..."

p17h0.75 change to "... 2. Design and initialize ...
The MDN of the state pdf ..."

p17h0.8 change to "... 3. The control is then initiated ...
..."

p17h0.9 change to "... 4. Approximate the state ...
..."

... and so on to step 10 ...
+----->>

p17h0.9 change to "... The dependency of these parameters on the MDN ..."

p20h0.33 change to "... recombined to generate the desired system response ..."

p20h0.95 change to "... state and control system values. This accounts for the system's uncertainty and improves ..."

++---------------------------++
C5. REFERENCES (using a quick web search, as opposed to checks using Scopus or standard indexes. I do not have access to CrossRefs “CrossCheck” via Elsevier's “iThenticate”)

C5a) Are references and citations in the standard format?
+-----+
Neural Networks journal example :
Chun, M., Biglou, J., Lenard, J., & Kim, J. (1999). Using neural networks to predict parameters in the hot working of aluminum alloys. Journal of Materials Processing Technology, vol. 86, pp. 245-251
Notes:
- Italicize the name of the publication (journal).
- Vol(number) can be as “9(10)”, for example.
- References should be sorted by the first author's last name.
- Citations in the text are of the form “... Chun, Biglou, Lenard, & Kim, 1999 ...” or “... Chun et al. 1999 ...” (not as reference #'s)
+-----+

Other than the numbering of the references in the list, the references are in the required format. (references - are not numbered for Neural Networks journal).

As per "C.5 Details" above:

p1h0.9 Citations : as a general comment, NN journal requires the format "... [Astrom 1970], [Werbos 1982] ..."
This applies to all citations.

C5b) Are references legitimate (using a quick web search and personal familiarity with references)?
By a spreadsheet random function, I randomly selected 5 references for checking :

[4] C. M. Bishop. Neural Networks for Pattern Recognition. Oxford University Press, New York, N.Y., 1995.
>> OK

[10] R. Herzallah and D. Lowe. A Bayesian perspective on stochastic neuro control. IEEE Transactions on Neural Networks, 19(5):914–924, May 2008.
>> OK

[12] Randa. Herzallah and Miroslav. K arna y. Fully probabilistic control design in an adaptive critic framework. Neural Networks, 24(11):1128–1135, 2011.
>> OK

[18] George G. Lendaris, Roberto A. Santiago, and Michael S. Carrol. Proposed framework for applying adaptive critics in real−time realm. In Proceedings of the 2002 International Joint Conference on Neural Networks, IJCNN'02, pages 1796–1801, Honolulu, HI , USA, 2002.
>> OK – but spell "Carroll"

[24] C.R. Rao. Linear method of statistical inference and their applications. Academia, Prague, 1987. in Czech.
>> OK

C5c) Is this paper significantly different from previous papers by the same authors? Is the work original?

It appears to be, but given the number of previous publications, I did not do a paper-by-paper check, and only briefly looked at a sampling of titles.

(This is the hardest reference check to do).

C5d) Is the relevant literature well represented in breadth and Depth?

Yes - I am familiar with a few of the references cited.

****

C6. LIMITATIONS OF THIS REVIEW

Reviewer's expertise on the subject: Low

I did not go back to the earlier papers, preferring to spend time on selected details in the current paper.

Although I read through the entire paper, and I was paying attention to the math, I only did a step-by-step check of the math of Theorem 1.

Bill@BillHowell.ca

****
C7. THOUGHTS: (again, changes to the paper are not require for these)
Here are some long-winded thoughts that are not really relevant to the paper review per se... For interest only, even if that.
These are separated from the "COMMENTS" above because they are less relevant to the actual paper.

Proofs of [existence, uniqueness, stability-convergence, optimality] - The author provides an interesting statement regarding optimality p11h0.4 :
“... The independence of ln(γ(xt−1)) on the optimised c(ut |xt−1) implies that the expression is minimised by the claimed pdf (13). ...”
But are [existence, uniqueness, stability-convergence] really guaranteed, upon which the “optimality” depends? Are they, too, provable?

Perhaps the author has been thinking of ways in which to “shape the modifiable distributions” (eg joint (s,c) pdf) to get different effects? (as opposed to Gaussian assumption)

p10h0.75 Equation (7) - An interesting side issue for me was the interpretation of the inverted-delta-square symbol, which I treated as a squared delta (as in gradient, as suggested by Equation (17) rather than a Laplacian (divergence of the gradient). However, I didn't actually DO anything with the interpretation other than simply reproduce the steps, and ran out of time before going into past papers. The reason that I am interested has nothing to do with this paper, but rather with attempts to replace Maxwell's equations in electromagnetism by resolving the [incomplete nature of the system of equations, inconsistent derivations, erroneous assumptions] built into the Heaviside 4-vector formulation, later overcome by Pointyng vecters relativity, and probably even quantum mechanics. In one of the effots, a triple-vector-crossProduct results, giving rise to a chiral term missing in electrodynamics. Like I said - no relation to the current paper, but it reminded me of another pre-occupation.

****
C8. MATH CHECKS - step-by-step

The following note may be ignored by the authors, as it is simply a record of this reviewer's step-by-step check over a part of the paper.

As a reviewer, I find that a step-by-step re-typing of a part of the paper as I have done below forces me to pay attention to details that I might otherwise skim over. Even though this is perhaps too time intensive to apply to the full paper, by doing so over part of the authors' work, it gives me far grater confidence in the rest of the paper, which is read, but not analysed step-by-step. It also gives the authors a better idea of the weaknesses of the reviewer!

+----+
Nomenclature :

A   denotes [scalar, vector, matrix]
A_T denotes transpose of A, also transpose(A
|A| denotes absolute value of matrix A (each element)
||A|| denotes spectral norm of A ||A||2
A_bar denotes an overscore on A
A_tilde denotes the authors' use of tilde over a Matrix symbol

******* denotes start/end of topics & sub-topics
+-----+ denotes steps in a [proof, development]..
denotes checks on specific steps by the reviewer (me) :
```
+----->> start Reviewer check :
<<-----+ end  Reviewer check :
```
>> short reviewer comments

Greek and Latin symbols are written in short text form.
These substitutions were necessary given the limited number of useful ASCII characters, the lack of superscript and subscript with simple text editors, and to make the text easier to use in software for expression processing (eventually - not ready yet at this time).

+----+

The step-by-step checks only covered up to the proof of Theorem 1, with significant checks starting with Equation (14) below.   I FELT THAT THE CENTRAL IMPORTANCE OF Theorem 1, made it more important to check in detail than the latter half of the paper (adaptive critic implementation and overall solution).

```
p2h0.5  the expected minimum cost–to–go function :
(1)    -ln(gamma(x(t-1)))
       = min(over c( u(t)|x(t-1) ) :
               integ( d(x(t),u(t)), indefinite :
                                   s(x(t)|x(t-1),u(t)) * c(u(t)|x(t-1))
                        * ( + ln(     s(x(t)|x(t-1),u(t)) * c(u(t)|x(t-1))
                                  / ( S(x(t)|x(t-1),u(t)) * C(u(t)|x(t-1)) )
                                 )
                            - ln(gamma(x(t))
                          )
                    )
```

)
.
and lowercase (s,c) represent the pdfs of the desired dynamics of the actual state vector and actual controller, respectively.
capitalcase (S,C) represent the pdfs of the desired dynamics of the observed state vector and ideal controller, respectively.
lowercase h is the most complete actual probabilistic description of the closed loop system, as factorised in (2) and (3) :

```
(2)   h( x(t),u(t-1)|x(t-1) ) = s( x(t)|u(t),x(t-1) ) * c( u(t)|x(t-1) )

(3)   H( x(t),u(t-1)|x(t-1) ) = S( x(t)|u(t),x(t-1) ) * C( u(t)|x(t-1) )
```

p3h0.15 For these stochastic systems, the pdf of optimal controller, c * (ut|xt-1),
minimising the cost–to–go function given in Equation (1) is shown to be generally determined
by the following functional recursion,

```
(4)   C_s( u(t)|x(t-1) )  =    C( u(t)|x(t-1) )
*exp[-beta( u(t),xh(t),x(t-1) ) ]
                             / gamma( x(t-1) )

      gamma( x(t-1) )     =  integ{ du(t), indefinite : C( u(t)|x(t-1) )
*exp[-beta( u(t),x(t-1) ) ] }

      beta( u(t),x(t-1) ) =  integ{ dx(t), indefinite :
                                              s( x(t)|u(t),x(t-1) )
                                    * ln[   s( x(t)|u(t),x(t-1) )
                                          / S( x(t)|u(t),x(t-1) )
                                        ]
                                  }
                             - ln(gamma(x(t)))
```
.
p5h0.75 Consider a stochastic discrete time nonlinear control system as follows

```
(5)   x(t) = f( x(t-1),u(t),epsilon(t) )
```

p6h0.35 For a general stochastic system of the form described in Equation (5), MDNs provide a
general framework for modelling the complete general pdf of the random state variables, s(xt |
xt-1 , ut) [4], [11]. Here, the pdf of the states, xt is represented as a linear combination of kernel
functions of the form

```
(6)     s( x(t)|x(t-1),u(t) ) =  sum[j=1 to M : alpha_j( x(t-1), u(t) ) *
f_j( x(t)|x(t-1),u(t) ) ]

(7)   f_j( x(t)|x(t-1),u(t) ) =  1 /   { (2*pi)^(n/2)*power[       sigma_j(t)
```

```
( x(t-1),u(t) ), n]  }
                                 * exp { * (-1/2)    *power[x(t) - x_j_hat(t)
( x(t-1),u(t) ), 2]
                                         /            power[      sigma_j(t)
( x(t-1),u(t) ), 2]
                                       }
```

p7h0.15 Measured process data, states and applied control inputs, is then used to train the MDN
and obtain the parameters of the state pdf. The MDN is then trained to minimise the negative
logarithm of the probability density function of the system states by using back–propagation

```
(8)   E = - sum(( over q : ln{ sum[j=1 to M : alpha_j( x(t-1),u(t) )
*f_j( x(t)|x(t-1),u(t) ) ] } ))
```

p8h0.0 the pdf of the control inputs can be approximated as a weighted sum of these control input kernel
functions as follows

```
(9)   c( u(t)|x(t-1) )   = sum[j=1 to M : phi_j(x(t-1))*g_j( u(t)|x(t-1) ) ]
```

p8h0.5 The kernel functions of the inverse controller are again taken to be Gaussian kernel functions,

```
(10)  g_j( u(t)|x(t-1) )     =  1 /
{ (2*pi)^(n/2)*power[            rho_j(t)( x(t-1),u(t) )   , r]  }
                                  * exp{ * (-1)      *power[ ||x(t) -
x_j_hat(t)( x(t-1),u(t) ) ||, 2]
                                         /
power[            rho_j(t)( x(t-1),u(t) )   , 2]
                                       }
```

p9h0.4 A method of interest in multimodal control applications is the most probable component from the MDN corresponding to the most probable branch. Taking the most probable branch from the system states pdf, the system states density function can be approximated by a Gaussian density as follows,

```
(11)  arg max (over j, alpha_j( x(t-1), u(t) )  ->  s( x(t)|x(t-1),u(t) )  =
f_j( x(t)|x(t-1),u(t) )
```

Similarly the control input pdf can be approximated by a Gaussian density as follows,

```
(12)  arg max (over j, phi_j(x(t-1))            ->  c( u(t)|x(t-1) )       =
g_j( u(t)|x(t-1) )
```

p9h0.75 Theorem 1 - the optimal controller minimising the cost–to–go function (1) is given by

```
(13)  Gs( u(t)|x(t-1) )          =    G( u(t)|x(t-1) )
```

```
*exp[-beta( u(t),xh(t),x(t-1) ) ]
                                       / gamma( x(t-1) )

      gamma( x(t-1) )              =  integ{ du(t), indefinite :
G( u(t)|x(t-1) )*exp[-beta( u(t),x_h(t),x(t-1) ) ] }

      beta( u(t),x_jh(t),x(t-1) ) =  integ{ dx(t), indefinite :
f_j( x(t)|u(t),x(t-1) )
                                                              *
ln[   f_j( x(t)|u(t),x(t-1) )
                                                                   /
F( x(t)|u(t),x(t-1) )
                                                                 ]
                                          }
                                     - ln(gamma_t(x(t)))


      - ln(gamma_t(x(t)))
      =  - ln(gamma(x_jh(t))) - 1/2*power((2 : Laplacian{ wrt x(t) :
ln(gamma(x(t)) |x_jh(t) * power[2 : sigma_j(t)( u(t), x(t-1) ) ] } ))

.
p10h0.2  combining (1), (7), (11), (12) using the chain rule :

(14)  - ln(gamma(x(t)))
      =  integ{ d(x(t),u(t)), indefinite :
                        f_j( x(t)|x(t-1),u(t) ) * g_j( u(t)|x(t-1) )
                  * [+ln((  f_j( x(t)|x(t-1),u(t) ) /
F( x(t)|x(t-1),u(t) )  ))
                     +ln((  g_j( u(t)|x(t-1)      ) /
G( u(t)|x(t-1)      )  ))
                     -ln(gamma(x(t))
                    ]
              }
.

+----->> start Reviewer check : combining (1), (11), (12)  to get (14), using
the chain rule

substituting into (1) the most probable f_j for s from (11), the most
probable g_j for c from (12)
                      the ideal         F       S          , the ideal
G       C
here I've "side-stepped" the chain rule, under the assumption that deriv( wrt
(x(t),u(t)) : f_j) = deriv( wrt (x(t),u(t)) : s) etc
    i.e. deriv( wrt s : f_j) = 1  etc
<1a>  -ln(gamma(x(t-1))) = min(over g_j( u(t)|x(t-1) ) :
  integ( d(x(t),u(t)), indefinite :
                        f_j(x(t)|x(t-1),u(t)) * g_j(u(t)|x(t-1))
               * [+ln((   f_j(x(t)|x(t-1),u(t)) * g_j(u(t)|x(t-1))
                       / (  F(x(t)|x(t-1),u(t)) *   G(u(t)|x(t-1)) )
                     ))
                  -ln(gamma(x(t))
```

```
                   ]
         )
   )
.
Rearranging
<1b>  -ln(gamma(x(t-1))) = min(over g_j( u(t)|x(t-1) ) :
  integ( d(x(t),u(t)), indefinite :
                          f_j(x(t)|x(t-1),u(t)) * g_j(u(t)|x(t-1))
              * [+ln((    f_j(x(t)|x(t-1),u(t)) /   F(x(t)|x(t-1),u(t)) ))
                 +ln((    g_j(u(t)|x(t-1)      ) /   G(u(t)|x(t-1)      ) ))
                 -ln(gamma(x(t))
                ]
       )
  )
.
. this is the same as (14), with the assumption that maximum probability f_j,
g_j ARE the minimum cost

<<-----+ end  Reviewer check : combining (1) and (11) , using the chain rule

p10h0.3 Now using Fubini theorem, Equation (14) can be recast as follows,

(15)  - ln(gamma(x(t)))
      =  integ{ du(t), indefinite :
                           g_j( u(t)|x(t-1) )
                * [        del( u(t),x_jh(t),x(t-1) )
                   + ln((  g_j( u(t)|x(t-1) )
                         /   G( u(t)|x(t-1) )
                       ))
                   - < ln(gamma(x(t)) >
                  ]
              }
.
where
(16)  del( u(t),x_jh(t),x(t-1) )
      =  integ{ dx(t), indefinite :
                          f_j( x(t)|u(t),x(t-1) )
                *   ln((  f_j( x(t)|u(t),x(t-1) )
                       /   F( x(t)|u(t),x(t-1) )
                      ))
              }
.

+----->> start Reviewer check : using Fubini theorem, Equation (14) can be
recast to (15)
reminder :  http://en.wikipedia.org/wiki/Fubini%27s_theorem
want to recast from d(x(t),u(t)) basis of (14)  to  [outer integral d(u(t))
and inner d(x(t))] basis of (15)

<14a> - ln(gamma(x(t)))  =
  integ{ d(u(t)), indefinite :
         integ{ d(x(t)), indefinite :
                          f_j( x(t)|x(t-1),u(t) ) * g_j( u(t)|x(t-1)      )
```

```
                * [+ln((  f_j( x(t)|x(t-1),u(t) ) /
F( x(t)|x(t-1),u(t) )  ))
                   +ln((  g_j( u(t)|x(t-1)      ) /
G( u(t)|x(t-1)      )  ))
                   -ln(gamma(x(t))
                  ]
              }
       }
.
```

ALL pdfs are conditional on x(t-1), but only some (those of f_j, F) are conditional on u(t)
Note that, once selected, the [f_j, F, g_j, G] are all FIXED Gaussian kernels for any iteration.
The question is, for the purposes of the integrations, can terms be considered SOLELY on the basis of their primary arguments?
For example, for the term f_j(x(t)|x(t-1),u(t)), can f_j be viewed as being solely a function of x(t), and not of u(t) ?

Here is a statement by the author on a different issue (joint s,c pdf), but perhaps reflecting partly some of the same issues :
p8h0.8 "... Finally the ideal joint pdf I h(xt , ut | x t-1) = I s(xt | xt-1 , ut) I c(ut | x t-1) is assumed to
be independent of the component kernels of the actual joint pdf h(xt , ut | xt-1) = s(xt |
xt-1 , ut)c(ut | xt-1). For example if a tracking control problem is considered, a reasonable form
for the ideal state pdf, I s(xt | xt-1 , ut) = If(xt | ut , xt-1) would be a Gaussian pdf with its mean
value being set to a desired value and variance to the non–reducible variance of innovations of
the active component. It specifies where the system state will be of high probability. ..."

A further assumption by the author SEEMS to be that gamma(x(t) should not be considered on the basis of x(t) alone. ??Why??

These assumptions seem to be a bit dangerous, but let's try it.

Separating terms:

```
<14b> -ln(gamma(x(t)))  =
  integ{ d(u(t)), indefinite :
         +  g_j( u(t)|x(t-1) )
                * [+ln((  g_j( u(t)|x(t-1)      ) /
G( u(t)|x(t-1)      )  ))
                   -  ln(gamma(x(t))
                  ]
                * integ{ d(x(t)), indefinite :
                          f_j( x(t)|x(t-1),u(t) )
                        }
         +  g_j( u(t)|x(t-1) )
                * integ{ d(x(t)), indefinite :
```

```
                          f_j( x(t)|x(t-1),u(t) ) *
ln((  f_j( x(t)|x(t-1),u(t) ) /   F( x(t)|x(t-1),u(t) )  ))
                       }
       }
.
let
<16a> <del> =
         integ{ d(x(t)), indefinite :
                          f_j( x(t)|x(t-1),u(t) ) *
ln((  f_j( x(t)|x(t-1),u(t) ) /   F( x(t)|x(t-1),u(t) )  ))
              }
. which is the expression for del in (16)

Also, the expression :
    integ{ d(x(t)), indefinite : f_j( x(t)|x(t-1),u(t) ) }
when integrated over the whole range of the pdf f_j, comes to 1.0

dropping the expression integ{ d(x(t)), indefinite :
f_j( x(t)|x(t-1),u(t) ) }, and substituting del into <14b>

<14c> -ln(gamma(x(t)))  =
  integ{ d(u(t)), indefinite :
         +  g_j( u(t)|x(t-1) )
                * [+ln((  g_j( u(t)|x(t-1)      ) /
G( u(t)|x(t-1)      )  ))
                   -  ln(gamma(x(t))
                  ]
         +  g_j( u(t)|x(t-1) )
                * del
       }
.
Rearranging :
<14d> -ln(gamma(x(t)))  =
  integ{ d(u(t)), indefinite :
         +  g_j( u(t)|x(t-1) )
                * [+del
                   +ln((  g_j( u(t)|x(t-1) ) /  G( u(t)|x(t-1) )  ))
                   -ln(gamma(x(t))
                  ]
       }
.
This is the same as Equation (15)


<<-----+ end  Reviewer check : using Fubini theorem, Equation (14) can be
recast to (15)


p10h0.7 the expected value of the cost–to–go function < ln(gamma(xt)) > can
be approximated by Taylor series to be given by

(17)  - < ln(gamma(x(t)) >
      =  -ln(gamma(x_j(t)) - 1/2*power((2 : Laplacian{ wrt x(t) :
```

```
ln(gamma(x(t)) |x_jh(t) * power[2 : sigma_j(t)( u(t), x(t-1) ) ] } ))
      =  -ln(gamma(x_j(t)) + 1/2*delta((wrt x(t) : lambda{ x(t))|x_jh(t) *
power[2 : sigma_j(t)( u(t), x(t-1) ) ] } ))

+----->>  Reviewer - here the index j designates the "winning" kernel ... as
per (11)  <<-----+

p10h0.85  Substitute Equation (17) into Equation (15) yield,

(18)  - ln(gamma(x(t)))
      =  integ{ du(t), indefinite :
                           g_j( u(t)|x(t-1) )
                * [+       del( u(t),x_jh(t),x(t-1) )
                   + ln((  g_j( u(t)|x(t-1) )
                         /   G( u(t)|x(t-1) )
                       ))
                   - ln(gamma(x_jh(t)) + 1/2*delta((wrt x(t) :
lambda{ x(t))|x_jh(t) * power[2 : sigma_j(t)( u(t), x(t-1) ) ] } ))
                  ]
              }
+----->>  no Reviewer check required for (18) - this is apparent  <<-----+

p11h0.0 Now introducing the following definition

(19)  beta( u(t),x_jh(t),x(t-1) )
      =                   del( u(t),x_jh(t),x(t-1) )
                   - ln(gamma(x_jh(t)) + 1/2*delta((wrt x(t) :
lambda{ x(t))|x_jh(t) * power[2 : sigma_j(t)( u(t), x(t-1) ) ] } ))
.

p11h0.1 Equation (18) can be rewritten as

(20)  - ln(gamma(x(t)))
      =  integ{ du(t), indefinite :
                           g_j( u(t)|x(t-1) )
                * [+      beta( u(t),x_jh(t),x(t-1) )
                   + ln((  g_j( u(t)|x(t-1) )
                         /   G( u(t)|x(t-1) )
                       ))
                  ]
              }
      =  integ{ du(t), indefinite :
                        g_j( u(t)|x(t-1) )
                * ln((  g_j( u(t)|x(t-1) )
                      / [ G( u(t)|x(t-1) )
*exp(-beta( u(t),x_jh(t),x(t-1) )) ]
                    ))
              }
      =  integ{ du(t), indefinite :
                              g_j( u(t)|x(t-1) )
                * ((  ln((  [ g_j( u(t)|x(t-1) ) * gamma(x(t-1))
]
                        / [ G( u(t)|x(t-1) )*exp(-beta( u(t),x_jh(t),x(t-1)
```

```
)) ]
                        ))
                    - ln(gamma(x(t-1)))
                  ))
              }
.

+----->> start Reviewer check : (18) to (20) using (19)

<18a> - ln(gamma(x(t)))
      =  integ{ du(t), indefinite :
                           g_j( u(t)|x(t-1) )
                * [+ ln((  g_j( u(t)|x(t-1) )
                         /   G( u(t)|x(t-1) )
                       ))
                   +       del( u(t),x_jh(t),x(t-1) )
                   - ln(gamma(x_jh(t)) + 1/2*delta((wrt x(t) :
lambda{ x(t))|x_jh(t) * power[2 : sigma_j(t)( u(t), x(t-1) ) ] } ))
                  ]
              }

<18b> - ln(gamma(x(t)))
      =  integ{ du(t), indefinite :
                           g_j( u(t)|x(t-1) )
                * [+ ln((  g_j( u(t)|x(t-1) )
                         /   G( u(t)|x(t-1) )
                       ))
                   + beta( u(t),x_jh(t),x(t-1) )
                  ]
              }
.same as first expression (20)

<18c> - ln(gamma(x(t)))
      =  integ{ du(t), indefinite :
                           g_j( u(t)|x(t-1) )
                * [+ ln((  g_j( u(t)|x(t-1) )
                         /   G( u(t)|x(t-1) )
                       ))
                   + ln(( exp((( beta( u(t),x_jh(t),x(t-1) ) ))) ))
                  ]
              }
.
<18d> - ln(gamma(x(t)))
      =  integ{ du(t), indefinite :
                           g_j( u(t)|x(t-1) )
                * [+ ln((  g_j( u(t)|x(t-1) )
                         /   G( u(t)|x(t-1) )
                         * exp((( beta( u(t),x_jh(t),x(t-1) ) )))
                       ))
                  ]
              }
.
<18e> - ln(gamma(x(t)))
```

```
      =  integ{ du(t), indefinite :
                        g_j( u(t)|x(t-1) )
                * ln((  g_j( u(t)|x(t-1) )
                      / [ G( u(t)|x(t-1) )*exp( -beta( u(t),x_jh(t),x(t-1) ))
]
                    ))
              }
. same as 2nd expression (20)
times & divide within ln(()) term by gamma(x(t-1)) :

<18f> - ln(gamma(x(t)))
      =  integ{ du(t), indefinite :
               +         g_j( u(t)|x(t-1) )
                 * ln((  g_j( u(t)|x(t-1) )
                       / [ G( u(t)|x(t-1) )*exp(
-beta( u(t),x_jh(t),x(t-1) )) ]
                       * gamma(x(t-1))
                       / gamma(x(t-1))
                     ))
              }

then move "/ gamma(x(t-1))" outside of (()) as "- ln(gamma(x(t-1)))", still
multiplierd by "g_j( u(t)|x(t-1) )"  :

<18g> - ln(gamma(x(t)))
      =  integ{ du(t), indefinite :
                             g_j( u(t)|x(t-1) )
                 * [ + ln((  g_j( u(t)|x(t-1) )*gamma(x(t-1))
                           / [ G( u(t)|x(t-1) )*exp(
-beta( u(t),x_jh(t),x(t-1) )) ]
                         ))
                     - ln(gamma(x(t-1)))
                   ]
              }
. this is the same as the 3rd equality in (20)

<<-----+  end Reviewer check : (18) to (20) using (19)
```

****

CONFIDENTIAL COMMENTS for review chair / committee use only:

None other than the list of Ratings

enddoc